FUNGAL BETA-GLUCURONIDASE GENES AND GENE PRODUCTS

REFERENCE TO SEQUENCE LISTING

[0001] The present invention includes a Sequence Listing submitted on compact disc, the contents of which are incorporated by reference in their entirety.

TECHNICAL FIELD

[0002] The present invention relates generally to β-glucuronidases, more specifically to β-glucuronidase derived from fungal species, and uses of these β-glucuronidases.

BACKGROUND OF THE INVENTION

[0003] The enzyme β-glucuronidase (GUS; E.C.3.2.1.31) hydrolyzes a wide variety of glucuronides. Virtually any aglycone conjugated to D-glucuronic acid through a β-O-glycosidic linkage is a substrate for GUS. In vertebrates, glucuronides containing endogenous as well as xenobiotic compounds are generated through a major detoxification pathway and excreted in urine and bile.

[0004] *Escherichia coli*, the major organism resident in the large intestine of vertebrates, utilizes the glucuronides generated in the liver and other organs as an efficient carbon source. In *E. coli*, β-glucuronidase is encoded by the gusA gene (Novel and Novel, *Mol. Gen. Genet. 120*:319-335, 1973), which is one member of an operon comprising two other protein-encoding genes: gusB encoding a permease (PER) specific for β-glucuronides, and gusC encoding an outer membrane protein (OMP) that facilitates access of glucuronides to the permease located in the inner membrane.

[0005] While β-glucuronidase activity is expressed in almost all tissues of vertebrates and their resident intestinal flora, GUS activity is absent in most other organisms. Notably, plants, many bacteria, and fungi have been reported to largely, if not completely, lack GUS activity. Thus, GUS is ideal as a reporter molecule in these organisms and has become the most widely used reporter system for plants.

[0006] In addition to use as a reporter molecule, GUS in combination with an innocuous glucuronide would be a preferred system to use for positive selection of transformed plants, especially for plants that will be consumed by humans. Because of the inefficiency of methods for transforming plant cells, only a small proportion of cells actually become transformed. Thus, it is desirable to select only those cells actually transformed. Typically, the selection methods involve transforming a cell with an antibiotic resistance gene along with the gene of interest and applying antibiotics to the cells, which kills the non-transformed cells.

[0007] Consumer resistance to antibiotic resistance genes has spurned research into alternative selection systems. Positive selection systems, wherein the transformed cells contain a gene whose gene product can utilize a compound that confers a growth advantage over the non-transformed cells. Ideally both the gene and the compound are biosafe to the environment and animals and humans.

[0008] GUS is the ideal system for positive selection for many reasons. First, biosafety assessment of GUS, including ecological and toxicological concerns, has shown GUS to be safe for both the environment and consumers (Gilissen et al. *Transgenic Res 7*: 157-163, 1998). Second, the gus gene is already present in several de-regulated food crops, such as papaya, beet and soybean, in the United States as well as in other countries. Third, the ease of making and isolating glucuronidated compounds allows a large choice of compounds to use for conferring growth advantage.

[0009] In positive selection systems under development, sugar compounds that plants do not normally metabolize, are being exploited in combination with xylose isomerase and mannose phosphate isomerase (U.S. Patent Nos. 5,994,629 and 5,767,378). Unfortunately, both of these systems have disadvantages: mannose is toxic to plant cells, some plants have endogenous xylose isomerase activity, and neither of the genes have undergone biosafety testing. Moreover, a reporter gene must still be used for visualization of transformed cells, a procedure that is necessary for confirmation of transformation. In addition, the intellectual property for these two systems is held by Syngenta who so far has not granted commercial licenses on terms favorable for small companies.

[0010] The *gus* gene in combination with a sugar glucuronide would provide the best positive selection system. GUS can serve as both a selectable and a reporter molecule; it is biosafe; and glucuronide sugars, such as cellobiuronic acid (a disaccharide comprising glucose and glucuronic acid) are readily isolated inexpensively. The *E. coli* *gus* gene, however, does not metabolize cellobiuronic acid. Therefore, there is a need for a GUS enzyme that can cleave cellobiuronic acid.

[0011] The present invention provides gene and protein sequences of fungal β-glucuronidases and variants thereof that are secreted and cleave cellobiuronic acid, while providing other related advantages.

SUMMARY OF THE INVENTION

[0012] In one aspect, an isolated nucleic acid molecule is provided comprising a nucleic acid sequence encoding a fungal β-glucuronidase. The fungus is a member of the Eurotiomycetes or Sordariomycetes class. On the basis of rRNA sequences, various isolates of fungus expressing β-glucuronidase are identified as members of *Penicillium*, *Eupenicillium, Scopulariopsis, Aspergillus,* or *Gibberella* (anamorph *Fusarium*) genera. In one embodiment nucleic acid sequences are provided for β-glucuronidases from *Penicillium canescens*, *Aspergillus nidulans*, *Scopulariopsis* sp., and *Gibberella zeae* (anamorph *Fusarium graminearum*). Further, the nucleic acid sequences encoding β-glucuronidases of *Penicillium canescens* and *Scopulariopsis* are provided both with and without sequence encoding a signal sequence, which directs proteins to rough endoplasmic reticulum. Certain embodiments provide for variants of the nucleic acid sequence, which vary in nucleotide sequence as a result of natural polymorphisms, site-directed mutagenesis, codon optimization and the like.

[0013] In other aspects, expression vectors comprising a gene encoding a fungal β-glucuronidase or a portion thereof that has enzymatic activity in operative linkage with a heterologous promoter are provided. In the expression vectors, the heterologous promoter may be selected from the group consisting of a developmental type-specific promoter, a tissue type-specific promoter, a cell type-specific promoter and an inducible promoter. The promoter should be functional in the host cell for the expression vector.

Examples of cell types include a plant cell, a bacterial cell, an animal cell and a fungal cell. In certain embodiments, the expression vector also comprises a nucleic acid sequence encoding a product of a gene of interest or portion thereof. The gene of interest may be under control of the same or a different promoter.

[0014] In other aspects, isolated fungal β-glucuronidase proteins are provided. Specific sequences are provided from *Penicillium*, *Eupenicillium*, *Scopulariopsis*, *Aspergillus*, or *Gibberella* (anamorph *Fusarium*) genera. In addition, β-glucuronidases from *Penicillium canescens* and *Scopulariopsis* are provided both with and without a signal sequence. Variants of the proteins are also provided. Methods to produce and purify the proteins of the present invention are described.

[0015] In another aspect, fusion proteins of a fungal β-glucuronidase or an enzymatically active portion thereof are provided. In certain embodiments, the fusion partner is a polypeptide chain of an antibody or fragment thereof that binds an antigen. Other fusion partners may be chosen to confer additional function or to facilitate purification of the β-glucuronidase protein.

[0016] The present invention also provides methods for monitoring expression of a gene of interest or a portion thereof in a host cell, comprising: (a) introducing into the host cell a vector construct, the vector construct comprising a nucleic acid molecule encoding a fungal β-glucuronidase of the present invention and a nucleic acid molecule encoding a product of the gene of interest or a portion thereof; (b) detecting the presence of the β-glucuronidase, thereby monitoring expression of the gene of interest. The fungal β-glucuronidases also have use in the present invention for confirming transformation of a host cell and for selecting transformed cells. In some preferred embodiments, the selecting compound is cellobiuronic acid, a disaccharide of glucose and glucuronic acid. In all these methods, a fungal glucuronide transport gene is optionally also introduced. These methods are especially useful in host cells that do not express an endogenous β-glucuronidase.

[0017] In another aspect, a method for providing an effector compound to a cell in a transgenic plant is provided. The method comprises (a) growing a transgenic plant that comprises an expression vector having a nucleic acid sequence encoding a fungal β-

glucuronidase in operative linkage with a heterologous promoter and a nucleic acid sequence comprising a gene encoding a cell surface receptor for an effector compound and (b) exposing the transgenic plant to a glucuronide, wherein the glucuronide is cleaved by the β-glucuronidase, such that the effector compound is released. This method is especially useful for directing glucuronides to particular and specific cells by further introducing into the transgenic plant a vector construct comprising a nucleic acid sequence that binds the effector compound. The effector compound can then be used to control expression of a gene of interest by linking a gene of interest with the nucleic acid sequence that binds the effector compound.

[0018] Transgenic plants and animals, such as aquatic animals and insects, that express a fungal β-glucuronidase are also provided. The present invention also provides seeds of transgenic plants.

[0019] These and other aspects of the present invention will become evident upon reference to the following detailed description and attached drawings. In addition, various references are set forth below which describe in more detail certain procedures or compositions (*e.g.*, plasmids, etc.), and are therefore incorporated by reference in their entirety.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] Figure 1 shows the amount of GUS enzyme activity in two fungal species at various times after addition of different inducers.

[0021] Figures 2A-C present the DNA sequence (SEQ ID NO:1) and the deduced amino acid sequence (SEQ ID NO:2) of the *gus* gene of *Scopulariopsis* sp. isolate RP38.3.

[0022] Figures 3A-C present the DNA sequence (SEQ ID NO:3) and the deduced amino acid sequence (SEQ ID NO:4) of the *gus* gene of *Penicillium canescens* isolate RPK.

[0023] Figures 4A-C present the DNA sequence (SEQ ID NO:5) and the deduced amino acid sequence (SEQ ID NO:6) of the *gus* gene of *Penicillium canescens* strain DSM 1215.

[0024] Figure 5 present the DNA sequence (SEQ ID NO:7) and the deduced amino acid sequence (SEQ ID NO:8) of the *gus* gene of *Gibberella zeae.*

[0025] Figure 6 present the DNA sequence (SEQ ID NO:9) and the deduced amino acid sequence (SEQ ID NO:10) of the *gus* gene of *Aspergillus nidulans..*

[0026] Figures 7A-E present alignments of amino acid sequences of GUS proteins from *C. elegans* (SEQ ID NO:11), *D. melanogaster* (SEQ ID NO:12), *M. musculus* (SEQ ID NO:13), *R. norvegicus* (SEQ ID NO:14), *F. catus* (SEQ ID NO:15), *C. familiaris* (SEQ ID NO:16), *C. aethiops* (SEQ ID NO:17), *H. sapiens* (SEQ ID NO:18), *S. solfataricus* (SEQ ID NO:19), *T. maritima* (SEQ ID NO:20), *L. gasseri* (SEQ ID NO:21), *E. coli* (SEQ ID NO:22), *Staphylococcus* sp. (SEQ ID NO:23), *A. nidulans* (SEQ ID NO:10), *P. canescens* (SEQ ID NO:4), *Scopulariopsis* sp. (SEQ ID NO:2), and *G. zeae* (SEQ ID NO:8).

[0027] Figure 8 is a schematic of pPWQ74.3, a vector backbone used to clone the *gus* genes of the present invention.

[0028] Figure 9 is a schematic of the vector pCAMBIA1305.2, the backbone of which was used to clone the *gus* genes of the present.

[0029] Figures 10A-B are pictographs of transgenic rice plants transformed with various constructs containing the *gus* genes of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

[0030] Prior to setting forth the invention, it may be helpful to an understanding thereof to set forth definitions of certain terms that will be used hereinafter.

[0031] As used herein, "β-glucuronidase" refers to an enzyme that catalyzes the hydrolysis of β-glucuronides. Assays and some exemplary substrates for determining β-glucuronidase activity, also referred to herein as GUS activity, are provided in U.S.

Patent No. 5,268,463. Other assays and substrates are taught in *GUS Protocols: Using the GUS gene as a reporter of gene expression* (ed. Gallagher SR, Academic Press, 1992, 221 pp.) In assays to detect β-glucuronidase activity, fluorogenic or chromogenic substrates are preferred. Such substrates include, but are not limited to, p-nitrophenyl β-D-glucuronide and 4-methylumbelliferyl β-D-glucuronide.

[0032] As used herein, the enzyme may be alternatively referred to as GUS or β-glucuronidase. The nucleic acid sequence that encodes GUS is referred to as *gus*. *gus* genes from particular species are written either as, for example, *E. coli gus* or preferably gus^{Eco}. If the *gus* gene is from an organism in which the genus is identified but the species is not, the superscript will use the first letters of the genus name.

[0033] As used herein, a "glucuronide" or "β-glucuronide" refers to an aglycone conjugated in a hemiacetal linkage, typically through the hydroxyl group, to the C1 of a free D-glucuronic acid in the β configuration. Glucuronides include, but are not limited to, O-glucuronides linked through an oxygen atom, S-glucuronides, linked through a sulfur atom, N-glucuronides, linked through a nitrogen atom and C-glucuronides, linked through a carbon atom (*see,* Dutton, *Glucuronidation of Drugs and Other Compounds,* CRC Press, Inc. Boca Raton, FL pp13-15). β-glucuronides consist of virtually any compound linked to the C1-position of glucuronic acid as a beta anomer, and are typically, though by no means exclusively, found as an O-glycoside. β-glucuronides are produced naturally in most vertebrates through the action of UDP-glucuronyl transferase as a part of the process of solubilizing, detoxifying, and mobilizing both natural and xenobiotic compounds, thus directing them to sites of excretion or activity through the circulatory system.

[0034] β-glucuronides in polysaccharide form are also common in nature, most abundantly in vertebrates, where they are major constituents of connective and lubricating tissues in polymeric form with other sugars such as N-acetylglucosamine (*e.g.*, chondroitin sulfate of cartilage, and hyaluronic acid, which is the principle constituent of synovial fluid and mucus). Other polysaccharide sources of β-glucuronides occur in bacterial cell walls, *e.g.,* cellobiuronic acid. β-glucuronides are relatively uncommon or absent in plants. Glucuronides and galacturonides found in

plant cell wall components (such as pectin) are generally in the alpha configuration, and are frequently substituted as the 4-O-methyl ether; hence, such glucuronides are not substrates for β-glucuronidase.

[0035] As used herein, a "variant" of *gus* or GUS is a nucleotide or amino acid sequence that contains one or more differences compared to the native sequence. Variants may arise naturally, e.g., polymorphisms, or be generated by *in vivo* or *in vitro* methods, a variety of these methods are described herein. Variants will have one or more amino acid or nucleotide alterations, one or more insertions, and/or one or more deletions.

[0036] As used herein, "percent sequence identity" is a percentage determined by the number of exact matches of amino acids or nucleotides to a reference sequence divided by the number of residues in the region of overlap. Within the context of this invention, preferred amino acid or nucleotide sequence identity for a variant of GUS is at least 75% and preferably greater than 80%, 85%, 90%, 95%, or 97%. Sequence identity may be determined by standard methodologies, including use of the National Center for Biotechnology Information BLAST search methodology available at www.ncbi.nlm.nih.gov. The identity methodologies preferred are non-gapped BLAST. However, those described in U.S. Patent 5,691,179 and Altschul *et al., Nucleic Acids Res.* 25:3389-3402, 1997, all of which are incorporated herein by reference, are also useful. Accordingly, if gapped BLAST 2.0 is utilized, then it is utilized with default settings.

[0037] As will be appreciated by those skilled in the art, a nucleotide sequence encoding fungal GUS may differ from wild-type sequences presented in the Figures, due to codon degeneracy, nucleotide polymorphisms, or amino acid differences. In certain embodiments, variants will hybridize to the wild-type nucleotide sequence at conditions of normal stringency, which is approximately 25-30°C below Tm of the native duplex (*e.g.*, 1 M Na+ at 65°C; *e.g.* 5X SSPE, 0.5% SDS, 5X Denhardt's solution, at 65°C or equivalent conditions; *see generally*, Sambrook *et al. Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Press, 1987; Ausubel *et al.*, *Current Protocols in Molecular Biology*, Greene Publishing, 1987). Alternatively,

the Tm can be calculated by the formula Tm=81.5 + 0.41%(G+C) – log[Na+]. Low stringency hybridizations are performed at conditions approximately 40°C below Tm, and high stringency hybridizations are performed at conditions approximately 10°C below Tm. Conditions suitable for hybridization of short nucleic acid molecules (less than about 500 bp) can be found in the references above. Note that some nucleic acid variants may not hybridize to the reference sequence because of codon degeneracy, such as degeneracy introduced for codon optimization in a particular host, in which case amino acid identity may be used to assess similarity of the variant to the native protein.

[0038] An "isolated nucleic acid molecule" refers to a polynucleotide molecule in the form of a separate fragment or as a component of a larger nucleic acid construct, that has been separated from its source cell (including the chromosome it normally resides in) at least once in a substantially pure form. Nucleic acid molecules may be comprised of a wide variety of nucleotides, including DNA, RNA, nucleotide analogues, have protein backbones (*e.g.,* PNA) or some combination of these. Similarly, an "isolated protein" refers to a protein that has been separated from its source cell.

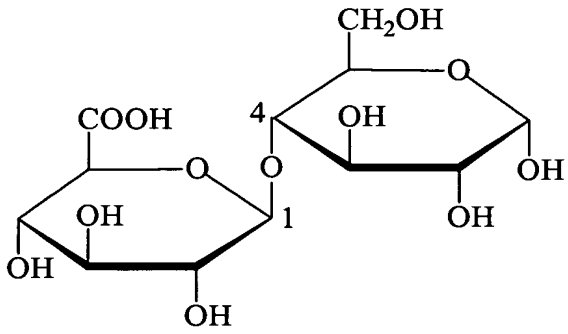
Fungal β-glucuronidase genes

[0039] As noted above, this invention provides gene sequences and gene products of fungal β-glucuronidases. As exemplified herein, genes from fungi, including the Eurotiomycetes and Sordariomycetes classes, that encode a β-glucuronidase are identified and characterized biochemically, genetically, and by DNA sequence analysis. Exemplary β-glucuronidase genes and their gene products from several genera, including *Penicillium*, *Scopulariopsis*, *Aspergillus*, and *Gibberella*, are provided herein. β-glucuronidase genes from additional fungi species may be identified as described herein or by hybridization of one of the fungal *gus* gene sequences to genomic or cDNA libraries, by genetic complementation, by function, by amplification, by antibody screening of an expression library and the like (*see* Sambrook *et al.*, *supra* Ausubel *et al.*, *supra* for methods and conditions appropriate for isolation of a β-glucuronidase from other species).

[0040] The presence of a fungal β-glucuronidase gene may be observed by a variety of methods and procedures. Particularly useful screens for identifying β-glucuronidase are biochemical screening for the gene product, genetic complementation, and sequence analysis comparisons.

[0041] Test samples containing fungi may be obtained from sources such as soil, plant surfaces, animal or human skin, decomposing matter, and the like. Fungal isolates are generally obtained by plating the sample (e.g., soil extract) on a suitable substrate in appropriate conditions. Conditions and substrates will vary according to the growth requirements of the fungus and the selecting compound. For example, when it is desirable to isolate fungi expressing a β-glucuronidase that cleaves cellobiuronic acid, samples are plated on minimal medium supplemented with vitamin and microelement solutions and with cellobiuronic acid as the sole carbon source.

[0042] Cellobiuronic acid (Cba) is the name by which the disaccharide having the following structure (I) is commonly known:



(I).

[0043] In the literature, the disaccharide of structure (I) is sometimes referred to by other names, including cellobiouronic acid, 4-*O*-(β-D-glucopyranuronosyl)-D-glucose, and β-glucuronosyl[1-4]glucose). *See, e.g.*, Carbohydrates, P.M. Collins, ed. Chapman and Hall, page 117, 1987. Regardless of the name, as shown in structure (I), cellobiuronic acid is a disaccharide formed between D-glucopyranuronic acid in β-linkage to a D-glucose, where the β-linkage is through carbon number 1 of D-glucopyranuronic acid and carbon number 4 of glucose (as identified in the structure (I)). A β linkage from a glucuronic acid to another sugar moiety (as seen in cellobiuronic acid) is referred to herein as a β-glucuronide linkage.

[0044] Other selective compounds can be used. Other saccharides or compounds required for growth of fungi that are in β linkage with a glucuronic acid may be used. Alternatively, the selecting molecule can be an S-glucuronide, linked through a sulfur atom, an N-glucuronide, linked through a nitrogen atom or a C-glucuronide, linked through a carbon atom to a saccharide or other compound required for cellular growth. Whatever the selecting glucuronide, fungi that express a β-glucuronidase may be identified by a glucuronide substrate that is readily detectable when cleaved by β-glucuronidase. If GUS enzymatic activity is present, the fungi will stain; a diffuse staining (halo) pattern surrounding a colony suggests that GUS is secreted.

[0045] The samples may contain bacteria or other microbes in addition to fungi. Some of these other microbes may have β-glucuronidase activity. Adhering bacteria or other microbes can be removed by consecutive sub-cultivation on medium containing antibiotics, such as ampicillin, streptomycin and nalidixic acid. Substrates such as deoxycholate, citrate, etc. may be used to inhibit other extraneous and undesired organisms such as gram-positive cocci and spore forming bacilli.

[0046] Following purification of the candidate fungi, it is prudent to verify GUS activity and cleavage of the selecting glucuronide by any of a number of different assays. In the Examples, the fungi were purified on YPD medium containing ampicillin, streptomycin and nalidixic acid and subsequently transferred back to the minimal medium containing Cba to reconfirm GUS activity by growth of the fungi. Alternatively, or in addition, a chromogenic assay for GUS activity can readily be performed by adding X-GlcA (5-bromo-4-chloro-3-indolyl-β-D-glucuronide) to the medium and observing whether a blue precipitate forms.

[0047] Other assays include *in vitro* biochemical assays, such as hydrolysis of a GUS substrate. Suitable GUS substrates are commercially available and widely known (*see,* U.S. Patent No: 5,268,463 and *GUS Protocols* (*supra*) for details of substrates and assays.] For example, hydrolysis of 4-methylumberlliferyl-β-D-glucuronide (MU-GlcA), a widely used GUS substrate, can be measured *in vitro*. For this assay, fungal isolates are grown and hyphal aggregates collected by e.g. vacuum filtration, washed and resuspended in minimal medium lacking glucuronides. Following a period of

starvation, various inducers of GUS activity (*e.g.* glucuronides) are added for an incubation time period. Aliquots of hyphal aggregates are collected at time intervals and proteins are extracted from these. The amount of cleavage of MU-GlcA by the test and control protein extracts are quantified, thereby confirming GUS activity.

[0048] A genetic complementation assay may be additionally performed to verify that the staining pattern is due to expression of a *gus* gene or to assist in isolating and cloning the *gus* gene. Briefly, in this assay, the candidate *gus* gene is transfected into an *E. coli* strain that is deleted for the *gus* operon (*e.g.*, KW1 described herein), and the staining pattern of the transfectant is compared to a mock-transfected host. Fungal genomic DNA, fungal cDNA, or an isolated *gus* gene is digested by *e.g.*, restriction enzyme reaction and ligated to a vector, which ideally is an expression vector. The recombinants are then transfected into a host strain, which preferably lacks or is deleted for any endogenous *gus* genes (*e.g.*, KW1 or a recA$^-$ deletion of KW1, called JEMA99.9). In some cases, the host strain may express the *gus* gene but preferably not in the compartment to be assayed. The transfected cells are selected on medium supplemented with an inducer of the *gus* gene. In the Examples, the fungal *gus* genes are cloned into a bacterial expression vector under control of the LAC promoter, expression of the *gus* gene is induced by IPTG (isopropyl-β-D-thiogalactoside), and β-glucuronidase activity is detected with X-GlcA. If GUS activity is present, the bacteria will turn blue; bacteria transfected with the vector alone will remain white. Moreover, if GUS is secreted, the transfectant should exhibit a diffuse staining pattern (halo) surrounding the colony.

[0049] The genera and species of the GUS-expressing fungi can be identified in myriad ways, including morphology, sequence similarity, metabolism signatures, and the like. A preferred method is comparison of rRNA sequence to sequences determined from known fungal genera or species. The rRNA sequences are generally obtained by sequencing of amplified fragments of genomic DNA. In fungal species, the 5.8S rRNA gene flanked by intergenic transcribed spacers 1 and 2 (ITS1, ITS2) have highly variable sequences and thus are well suited for identification of fungi. Preferably the match is at least 90%, at least 95%, or at least 99%. If no perfect match, or near perfect match,

with a known species is found or if additional confirmation is desirable, sequence obtained of the 18S rRNA gene is compared to a database of fungal 18S rRNA sequences to establish the phylogenetic placement at the genus level. Nucleotide identity is preferably at least 90%, at least 95%, or at least 99%. For either of these rRNA sequences, a suitable method to obtain sequence is to amplify the genes using primers that derive from conserved regions and subject the amplified fragments to DNA sequence analysis. Other methods to isolated and determine sequence rRNA gene regions are well known. Occasionally fungal species represented in the databases may be renamed or reclassified in a different genus. In such cases, other of these fungi, which are isolated and characterized, such as those herein, will also change accordingly.

[0050] In exemplary screens, three isolates of fungi that can utilize Cba as a carbon source and have GUS activity are obtained from soil samples. Confirmation of GUS activity is established by biochemical assay and growth of purified fungi on medium containing Cba. rRNA sequence analyses and comparison to other eukaryotic rRNA genes identified the fungi as *Penicillium canescens* and *Scopulariopsis* sp.

[0051] The fungal *gus* gene can be isolated by any number of methods. For example, it can be cloned by inserting genomic DNA or cDNA fragments into an expression vector and looking for complementation in a *gus* deletion strain. The vector with the insert is then recovered by isolation or the insert is amplified and recovered. Another method is to amplify the *gus* gene from genomic DNA or cDNA using primers derived from conserved areas of known *gus* genes from bacteria and animals. In the Examples, a 1.2 kb signature fragment of the *gus* gene is amplified from fungal DNA from the three isolates. The complete nucleotide sequences of the *gus* genes, including upstream and downstream non-coding sequences are obtained by amplification, but could be isolated in other ways such as using the 1.2 kb fragment as a probe against a genomic library or a cDNA library. Other well-known methods can alternatively be used.

[0052] DNA sequences of the *gus* gene contained in these three isolates are presented in Figures 2-4 and as SEQ ID NOs:1, 3, and 5. Translation of a continuous open reading frame reveals a 641 amino acid (*Scopulariopsis*) protein and a 634 amino acid protein (*P. canescens*). Furthermore, there appears to be signal peptides with predicted

cleavage positions at amino acids 26-27 (*Scopulariopsis*) and 18-19 (*P. canescens*), which would then yield mature proteins of 615 and 616 amino acids, respectively.

[0053] Confirmation that the ORFs encode β-glucuronidases is made by sequence similarity between the predicted fungal protein sequences and bacterial and animal GUS protein sequences. As demonstrated herein, there is significant similarity to microbial and mammalian β-glucuronidases. Furthermore, it is confirmed that conserved domains and signature sequences common to family 2 glycosyl hydrolases (e.g., β-glucuronidase) are present in fungal β-glucuronidases (Figures 7A-D). The amino acid sequences are shown in alignment in Figures 7A-D. The signature peptide sequences for family 2 glycosyl hydrolases (Henrissat, *Biochem Soc Trans 26:*153, 1998; Henrissat B *et al., FEBS Lett 27:*425, 1998) are located from amino acids 423 to 448 and from amino acids 498 to 512 (consensus numbering in Figures 7A-D). The acid/base catalyst is Glu 512 (consensus numbering) and the catalytic nucleophile (proton donor) is Glu 608 (Wong *et al., J. Biol Chem. 18*: 34057, 1998). Overall identity (similarity) between *Scopulariopsis* and *E. coli* GUS proteins is 49.6% (60.5%), between *Penicillium* and *E. coli* is 50.3% (61.6%). Identity at the DNA level is 55.3% (between *Scopulariopsis* and *E. coli*) and 50.8% (between *Penicillium* and *E. coli*).

[0054] There are four Asn-Xaa-Ser/Thr sequences in *Penicillium* and five Asn-Xaa-Ser/Thr sequences in *Scopulariopsis* that may serve as site for N-glycosylation in the ER. Furthermore, unlike the *E. coli* and human β-glucuronidases, which have 9 and 4 cysteines respectively, these GUS proteins have two Cys residues.

[0055] Additional fungi that have a *gus* gene can be identified by any of the methods described herein or by interrogation of sequences in a database. In the Examples, two additional *gus* genes are identified in a publicly available dataset. The *gus* genes are found in *Aspergillus nidulans* and *Gibberella zeae*. The *gus* gene sequences from these species and from other fungal species can be isolated as described herein, e.g., amplification using primers derived from conserved regions or from sequences of the genes as published in a database, by hybridization of genomic or cDNA libraries with a known *gus* sequence, and the like.

[0056] In certain aspects, the present invention provides fragments of fungal *gus* genes. A fragment is any length sequence. Fragments of fungal *gus* may be isolated or constructed for use in the present invention. For example, restriction fragments can be isolated by well-known techniques from template DNA, *e.g.*, plasmid DNA, and DNA fragments, including, but not limited to, digestion with restriction enzymes or amplification. These fragments may be used in hybridization methods (*see,* exemplary conditions described *infra*) or inserted into an appropriate vector for expression or production. In other embodiments, oligonucleotides (two or more nucleotides) of fungal GUSes are provided especially for use as amplification primers. In such case, the oligonucleotides are at least 12 bases and preferably at least 15 bases (*e.g.,* at least 18, 21, 25, 30 bases) and generally not longer than 50 bases. It will be appreciated that any of these fragments described herein can be double-stranded, single-stranded, derived from coding strand or complementary strand and be exact or mismatched sequence.

[0057] Other fragments (oligonucleotides) for use in this invention may be at least 12 nucleotides long (*e.g.,* at least 15 nt, 17 nt, 20 nt, 25 nt, 30 nt, 40 nt, 50 nt, 100 nt, 150 nt, 200 nt and so on). One skilled in the art will appreciate that other methods are available to obtain DNA or RNA molecules having at least a portion of a fungal *gus* sequence. Other uses for fragments include hybridization and isolation of new fungal *gus* genes, amplification, site-directed mutagenesis and the like. Moreover, for particular applications, these nucleic acids may be labeled by techniques known in the art, such as with a radiolabel (*e.g.*, ^{32}P, ^{33}P, ^{35}S, ^{125}I, ^{131}I, ^{3}H, ^{14}C), fluorescent label (*e.g.*, FITC, Cy5, RITC, Texas Red), chemiluminescent label, enzyme, biotin and the like.

[0058] In certain aspects, the fragments have sequences of one or both of the signatures or have sequence from at least some of the more highly conserved regions of GUS (*e.g.,* from approximately amino acids 423 to 448 and from amino acids 498 to 512 based on the consensus numbering in Figure 7A-E). In the various embodiments, useful fragments comprise those nucleic acid sequences which encode at least the glutamate residue that acts as the acid/base catalyst (amino acid position 512) and the

glutamate residue that acts as the catalytic nucleophile at position 608 (consensus numbering in Figure 7A-E).

Fungal β-glucuronidase gene products

[0059] The present invention also provides β-glucuronidase gene products in various forms. Forms of GUS protein include, but are not limited to, secreted forms, membrane-bound forms, cytoplasmic forms, fusion proteins, chemical conjugates of GUS and another molecule, portions of GUS protein, and other variants. GUS protein may be produced by expression from a recombinant vector, biochemical isolation from natural sources such as hyphae, from transformed host cells, and the like.

[0060] In certain aspects, variants of secreted fungal GUS are useful within the context of this invention. Variants include nucleotide or amino acid substitutions, deletions, insertions, and chimeras (*e.g.,* fusion proteins). Typically, when the result of synthesis, amino acid substitutions are conservative, *i.e.*, substitution of amino acids within groups of polar, non-polar, aromatic, charged, etc. amino acids.

[0061] Variants may be constructed by any of the well known methods in the art (*see, generally,* Ausubel *et al.*, *supra*; Sambrook *et al.*, *supra*). Such methods include site-directed oligonucleotide mutagenesis, restriction enzyme digestion and removal or insertion of bases, amplification using primers containing mismatches or additional nucleotides, splicing of another gene sequence to the native fungal *gus* gene, synthesis and the like. Briefly, preferred methods for generating a few nucleotide substitutions utilize an oligonucleotide that spans the base or bases to be mutated and contains the mutated base or bases. The oligonucleotide is hybridized to complementary single stranded nucleic acid and second strand synthesis is primed from the oligonucleotide. Similarly, deletions and/or insertions may be constructed by any of a variety of known methods. For example, the gene can be digested with restriction enzymes and religated such that some sequence is deleted or ligated with an isolated fragment having cohesive ends so that an insertion or large substitution is made. In other embodiments, variants are generated by shuffling of regions (see U.S. Patent No. 5,605,793) or by "molecular evolution" techniques (see U. S. Patent No. 5,723,323). Other means to generate

variant sequences may be found, for example, in Sambrook *et al.* (*supra*) and Ausubel *et al.* (*supra*).

[0062] In addition to directed mutagenesis in which one or a few amino acids are altered, variants that have multiple substitutions may be generated. The substitutions may be scattered throughout the protein or functional domain or concentrated in a small region. For example, a region may be mutagenized by oligonucleotide-directed mutagenesis in which the oligonucleotide contains a string of dN bases or the region is excised and replaced by a string of dN bases. Thus, a population of variants with a randomized amino acid sequence in a region is generated. The variant with the desired properties (*e.g.*, more efficient secretion) is then selected from the population.

[0063] Verification of variant sequences is typically accomplished by restriction enzyme mapping, sequence analysis, and/or probe hybridization, although other methods may be used. The double-stranded nucleic acid is transformed into host cells, typically *E. coli*, but alternatively, other prokaryotes, yeast, or larger eukaryotes may be used. Standard screening protocols, such as nucleic acid hybridization, amplification, and DNA sequence analysis, can be used to identify mutant sequences.

[0064] In preferred embodiments, the protein and variants are capable of being secreted and cleaving Cba. A GUS protein is secreted if the amount of secretion expressed as a secretion index is statistically significantly higher for the candidate protein compared to a standard, typically *E. coli* GUS. The secretion index may be calculated as the percentage of total GUS activity in periplasm or other extracellular environment less the percentage of total β-glucuronidase activity found in the same extracellular environment for a non-secreted GUS. Cleavage of Cba can be determined *in vitro*, e.g., by thin layer chromatography, or *in vivo*, e.g., survival of transformed cells on Cba as sole carbon source.

[0065] In other embodiments, variants may be directed to other cellular compartments, such as membrane or cytoplasm. Membrane-spanning amino acid sequences are generally hydrophobic and many examples of such sequences are well-known. These sequences may be spliced onto fungal secreted GUS by a variety of methods including conventional recombinant DNA techniques. Similarly, sequences that direct proteins to

cytoplasm (*e.g.*, Lys-Asp-Glu-Leu) may be added to the reference GUS, typically by recombinant DNA techniques.

[0066] In other embodiments, variants of fungal GUS are capable of binding to a hapten, such as biotin, dinitrophenol, and the like. Binding assays to such haptens are well known and may be found, for example, in *Antibodies: A Laboratory Manual* (*infra*)

[0067] In other embodiments, a fusion protein comprising GUS may be constructed from the nucleic acid molecule encoding fungal *gus* and one or more other nucleic acid molecules. As will be appreciated, the fusion partner gene may contribute, within certain embodiments, an open reading frame. In preferred embodiments, fungal GUS is fused to avidin, streptavidin or one of the polypeptides of an antibody. Thus, it may be desirable to use only the catalytic region of GUS (*e.g.*, the region containing the two well-defined catalytically active amino acid residues plus optionally the conserved family 2 signatures). The choice of the fusion partner depends in part upon the desired application. The fusion partner may be used to alter specificity of GUS, provide a reporter function, provide a tag sequence for identification or purification protocols, and the like. The reporter or tag can be any protein or peptide that allows convenient and sensitive measurement or facilitates isolation of the gene product and does not interfere with the function of GUS. For example, green fluorescent protein and β-galactosidase are readily available as DNA sequences and may be used to provide additional function to GUS. A peptide tag is a short sequence, usually derived from a native protein, which is recognized by an antibody, hapten, or other molecule. Peptide tags include, but are not limited to, FLAG®, Glu-Glu tag (Chiron Corp., Emeryville, CA), KT3 tag (Chiron Corp.), T7 gene 10 tag (Invitrogen, La Jolla, CA), T7 major capsid protein tag (Novagen, Madison, WI), His_6 (hexa-His), and HSV tag (Novagen). Besides these tags, other proteins or peptides, such as glutathione-S-transferase may be used as a tag.

[0068] In other aspects of the present invention, isolated fungal glucuronidase proteins are provided. In one embodiment, GUS protein is expressed as a hexa-His fusion protein and isolated by metal-affinity chromatography, for example using nickel-coupled beads. Briefly, a sequence encoding His_6 is linked to a DNA sequence

encoding a GUS. Although the His_6 sequence can be positioned anywhere in the molecule, it is typically linked at the 3' end immediately preceding the termination codon. The hexa-His-GUS fusion may be constructed by any of a variety of methods. A convenient method is amplification of the *gus* gene using a downstream primer that contains the codons for His_6. Alternatively, the *gus* gene may be cloned into a vector that already contains the His_6 coding sequence.

[0069] Alternatively, β-glucuronidase protein, with or without a tag, may be isolated by standard methods, such as affinity chromatography using matrices containing saccharo-lactone, phenyl-thio-β-glucuronide, antibodies to GUS protein and the like, size exclusion chromatography, ionic exchange chromatography, HPLC, and other known protein isolation methods (*see generally* Ausubel *et al. supra*; Sambrook *et al. supra*). The protein can be expressed as a hexa-His fusion protein and isolated by metal-affinity chromatography, for example with nickel-coupled beads. An isolated purified protein gives a single band on SDS-PAGE when stained with Coomassie brilliant blue.

[0070] In one aspect of the present invention, peptides having fungal GUS sequence are provided. Peptides may be used as immunogens to raise antibodies, as well as other uses, such as competitive inhibitors in assays. Peptides are generally five to 100 amino acids long, and more usually 10 to 50 amino acids. Peptides are readily chemically synthesized in an automated fashion (*e.g.*, PerkinElmer, ABI Peptide Synthesizer) or may be obtained commercially. Peptides may be further purified by a variety of methods, including high-performance liquid chromatography (HPLC). Furthermore, peptides and proteins may contain amino acids other than the 20 naturally occurring amino acids or may contain derivatives and modification of the amino acids.

Antibodies to fungal GUS

[0071] Antibodies to fungal GUS proteins, fragments, or peptides discussed herein may readily be prepared. Such antibodies may specifically recognize reference fungal GUS protein and not a variant protein, or variant protein and not wild type protein, or equally recognize both the mutant (or variant) and wild-type forms. Antibodies may be

used for isolation of the protein, inhibiting activity of the protein (antagonist), or enhancing activity of the protein (agonist).

[0072] Within the context of the present invention, antibodies are understood to include monoclonal antibodies, polyclonal antibodies, anti-idiotypic antibodies, antibody fragments (*e.g.*, Fab, and $F(ab')_2$, F_V variable regions, or complementarity determining regions). Antibodies are generally accepted as specific against GUS protein if they bind with a K_d of greater than or equal to 10^{-7} M, preferably greater than of equal to 10^{-8} M. The affinity of a monoclonal antibody or binding partner can be readily determined by one of ordinary skill in the art (*see* Scatchard, *Ann. N.Y. Acad. Sci. 51*:660-672, 1949).

[0073] Briefly, a polyclonal antibody preparation may be readily generated in a variety of warm-blooded animals such as rabbits, mice, or rats by well-known procedures. Monoclonal antibodies may be readily generated from hybridoma cell lines using conventional techniques (*see* U.S. Patent Nos. RE 32,011, 4,902,614, 4,543,439, and 4,411,993; *see also Antibodies: A Laboratory Manual*, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988). Other techniques may also be utilized to construct monoclonal antibodies (*see* Huse *et al.*, *Science 246*:1275-1281, 1989; Sastry *et al.*, *Proc. Natl. Acad. Sci. USA 86*:5728-5732, 1989; Alting-Mees *et al.*, *Strategies in Molecular Biology 3*:1-9, 1990; describing recombinant techniques).

[0074] One of ordinary skill in the art will appreciate that a variety of alternative techniques for generating antibodies exist. In this regard, the following U.S. patents teach a variety of these methodologies and are thus incorporated herein by reference: U.S. Patent Nos. 5,840,479; 5,770,380; 5,204,244; 5,482,856; 5,849,288; 5,780,225; 5,395,750; 5,225,539; 5,110,833; 5,693,762; 5,693,761; 5,693,762; 5,698,435; and 5,328,834.

[0075] Once suitable antibodies have been obtained, they may be isolated or purified by many techniques well known to those of ordinary skill in the art (*see Antibodies: A Laboratory Manual*, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988). Suitable techniques include peptide or protein affinity columns, HPLC (*e.g.*,

reversed phase, size exclusion, ion-exchange), purification on protein A or protein G columns, or any combination of these techniques.

Assays for function of β-glucuronidase

[0076] In preferred embodiments, fungal β-glucuronidases will at least have enzymatic activity and in other preferred embodiments, will also have the capability of being secreted. As noted above, variants of these reference GUS proteins may exhibit altered functional activity and cellular localization. Enzymatic activity may be assessed by assays such as the ones disclosed herein or in U.S. Patent No. 5,268,463 (Jefferson). Generally, a chromogenic or fluorogenic substrate is incubated with cell extracts, tissue or tissue sections, or purified protein. Cleavage of the substrate is monitored by a method appropriate for the aglycone or the glucuronic acid that is released.

[0077] A variety of methods may be used to demonstrate that a β-glucuronidase is secreted. For example, a rapid screening method in which colonies of organisms or cells, such as bacteria, yeast or insect cells, are plated and incubated with a readily visualized glucuronide substrate, such as X-GlcA. A colony with a diffuse staining pattern likely secretes GUS, although such a pattern could indicate that the cell has the ability to pump out the aglycone or its dimer, that the cell has become leaky, or that the enzyme is membrane bound. These unlikely alternatives can be ruled out by using a host cell for transfection that does not pump out the aglycone chosen and lacks an endogenous *gus* gene.

[0078] Secretion of the enzyme may be verified by assaying for GUS activity in the extracellular environment. If the cells secreting GUS are gram-positive bacteria, yeasts, molds, plants, or other organisms with cell walls, activity may be assayed in the culture medium and in a cell extract, however, the protein may not be transported through the cell wall. Thus, if no or low activity of a secreted form of GUS is found in the culture medium, protoplasts are prepared by osmotic shock or enzymatic digestion of the cell wall or any other suitable procedure, and the supernatant is assayed for GUS activity. If the cells secreting GUS are gram-negative bacteria, the culture supernatant is tested, but β-glucuronidase may be retained in the periplasmic space between the inner and outer

membrane. In this case, spheroplasts are prepared by osmotic shock, enzymatic digestion, or any other suitable procedure, and the supernatant is assayed for GUS activity. Cells without cell walls are assayed for GUS in cell supernatant and cell extract. The fraction of activity in each compartment is compared to the activity of a non-secreted GUS in the same or similar host cells. A β-glucuronidase is secreted if significantly more enzyme activity than *E. coli* GUS activity is found in extracellular spaces. The amount of secretion is generally normalized to the amount of a non-secreted protein (e.g., *β*-galactosidase) found in intracellular spaces. By this assay, usually less than 10% of *E. coli* GUS is secreted. Within the context of this invention, higher amounts of secreted enzyme are preferred (*e.g.*, greater than 20%, 25%, 30%, 40%, 50%).

[0079] β-glucuronidases that exhibit particular substrate specificity are also useful within the context of the present invention. As noted above, glucuronides can be linked through an oxygen, carbon, nitrogen or sulfur atom. Glucuronide substrates having each of the linkages may be used in one of the assays described herein to identify GUSes that discriminate among the linkages. In addition, various glucuronides containing a variety of aglycones may be used to identify GUSes that discriminate among the aglycones.

Vectors, host cells and means of expressing and producing protein

[0080] Fungal β-glucuronidase may be expressed in a variety of host organisms. For protein production and purification, GUS is preferably secreted and produced in bacteria, such as *E. coli*, for which many expression vectors have been developed and are available. Other suitable host organisms include other bacterial species (*e.g., Bacillus*), and eukaryotes, such as yeast (*e.g., Saccharomyces cerevisiae*), mammalian cells (*e.g.*, CHO and COS-7), plant cells and insect cells (*e.g.*, Sf9). Vectors for these hosts are well known.

[0081] A DNA sequence encoding a fungal β-glucuronidase is introduced into an expression vector appropriate for the host. The sequence is derived from an existing clone or synthesized. As described herein, a fragment of the coding region may be used,

but if enzyme activity is desired, the catalytic region should be included. A preferred means of synthesis is amplification of the gene from cDNA, genomic DNA, or a recombinant clone using a set of primers that flank the coding region or the desired portion of the protein. Restriction sites are typically incorporated into the primer sequences and are chosen with regard to the cloning site of the vector. If necessary, translational initiation and termination codons can be engineered into the primer sequences. The sequence of GUS can be codon-optimized for expression in a particular host. For example, a secreted form of β-glucuronidase isolated from a bacterial species that is expressed in a fungal host, such as yeast, can be altered in nucleotide sequence to use codons preferred in yeast. Codon-optimization may be accomplished by methods such as splice overlap extension, site-directed mutagenesis, automated synthesis, and the like.

[0082] At minimum, an expression vector must contain a promoter sequence. Other regulatory sequences may be included. Such sequences include a transcription termination signal sequence, secretion signal sequence, intron, enhancer, origin of replication, selectable marker, and the like. The regulatory sequences are operationally associated with one another to allow transcription or translation.

[0083] Suitable host cells may be prokaryotic or eukaryotic. The most commonly used bacteria is *E. coli*, but any transformable bacteria may alternatively be used. Eukaryotic cells useful in this invention include, but are not limited to, yeast cells, plant cells, mouse cells, and human cells. A host cell may be cells that grow as isolated cells or may be an organized collection of cells, such as meristem tissue, callus tissue or other explanted tissue from plants. Human organisms are specifically excluded from host cells, although isolated human cells may be used.

Expression in bacteria

[0084] The plasmids used herein for expression of secreted GUS include a promoter designed for expression of the proteins in a bacterial host. Suitable promoters are widely available and are well known in the art. Inducible or constitutive promoters are preferred. Such promoters for expression in bacteria include promoters from the T7

phage and other phages, such as T3, T5, and SP6, and the trp, lpp, and lac operons. Hybrid promoters (*see,* U.S. Patent No. 4,551,433), such as tac and trc, may also be used. Promoters for expression in eukaryotic cells include the P10 or polyhedron gene promoter of baculovirus/insect cell expression systems (*see, e.g.*, U.S. Patent Nos. 5,243,041, 5,242,687, 5,266,317, 4,745,051, and 5,169,784), MMTV LTR, RSV LTR, SV40, metallothionein promoter (*see, e.g.*, U.S. Patent No. 4,870,009) and other inducible promoters. For protein expression, a promoter is inserted in operative linkage with the coding region for β-glucuronidase.

[0085] The promoter controlling transcription of β-glucuronidase may be controlled by a repressor. In some systems, the promoter can be de-repressed by altering the physiological conditions of the cell, for example, by the addition of a molecule that competitively binds the repressor, or by altering the temperature of the growth media. Preferred repressor proteins include, but are not limited to the *E. coli* LACI repressor responsive to IPTG induction, the temperature sensitive λcI857 repressor, and the like. The *E. coli* LACI repressor is preferred.

[0086] In other preferred embodiments, the vector also includes a transcription terminator sequence. A "transcription terminator region" has either a sequence that provides a signal that terminates transcription by the polymerase that recognizes the selected promoter and/or a signal sequence for polyadenylation.

[0087] Preferably, the vector is capable of replication in host cells. Thus, for bacterial hosts, the vector preferably contains a bacterial origin of replication. Preferred bacterial origins of replication include the f1-ori and col E1 origins of replication, especially the origin derived from pUC plasmids.

[0088] The plasmids also preferably include at least one selectable gene that is functional in the host. A selectable gene includes any gene that confers a phenotype on the host that allows transformed cells to be identified and selectively grown. Suitable selectable marker genes for bacterial hosts include the ampicillin resistance gene (Amp^r), tetracycline resistance gene (Tc^r) and kanamycin resistance gene (Kan^r). Suitable markers for eukaryotes usually complement a deficiency in the host (*e.g.*,

thymidine kinase (tk) in tk- hosts). However, drug markers are also available (*e.g.*, G418 resistance and hygromycin resistance).

[0089] The sequence of nucleotides encoding β-glucuronidase may also include a classical secretion signal, whereby the resulting peptide is a precursor protein processed and secreted. The resulting processed protein may be recovered from the periplasmic space or the fermentation medium. Secretion signals suitable for use are widely available and are well known in the art (von Heijne, *J. Mol. Biol. 184*:99-105, 1985). Prokaryotic and eukaryotic secretion signals that are functional in *E. coli* (or other host) may be employed. The presently preferred secretion signals include, but are not limited to pelB, matα, extensin and glycine-rich protein.

[0090] One skilled in the art appreciates that there are a wide variety of suitable vectors for expression in bacterial cells and which are readily obtainable. Vectors such as the pET series (Novagen, Madison, WI) and the tac and trc series (Pharmacia, Uppsala, Sweden) are suitable for expression of a β-glucuronidase. A suitable plasmid is ampicillin resistant, has a colEI origin of replication, a $lacI^q$ gene, a lac/trp hybrid promoter in front of the lac Shine-Dalgarno sequence, a hexa-his coding sequence that joins to the 3' end of the inserted gene, and an rrnB terminator sequence.

[0091] The choice of a bacterial host for the expression of a β-glucuronidase is dictated in part by the vector. Commercially available vectors are paired with suitable hosts. The vector is introduced in bacterial cells by standard methodology. Typically, bacterial cells are treated to allow uptake of DNA (for protocols, see generally, Ausubel *et al.*, *supra*; Sambrook *et al.*, *supra*). Alternatively, the vector may be introduced by electroporation, phage infection, or another suitable method.

Expression in plant cells

[0092] As noted above, the present invention provides vectors capable of expressing fungal secreted β-glucuronidase and secreted fungal β-glucuronidases. For agricultural applications, the vectors should be functional in plant cells. Suitable plants include, but are not limited to, wheat, rice, corn, soybeans, lupins, vegetables, potatoes, canola, nut

trees, coffee, cassava, yam, alfalfa and other forage plants, cereals, legumes and the like. In one embodiment, rice is a host for *gus* gene expression.

[0093] Vectors that are functional in plants are preferably binary plasmids derived from *Agrobacterium* plasmids. Such vectors are capable of transforming plant cells. These vectors contain left and right border sequences that are required for integration into the host (plant) chromosome. At minimum, between these border sequences is the gene to be expressed under control of a promoter. In preferred embodiments, a selectable gene is also included. The vector also preferably contains a bacterial origin of replication for propagation in bacteria.

[0094] A gene for fungal β-glucuronidase should be in operative linkage with a promoter that is functional in a plant cell. Typically, the promoter is derived from a host plant gene, but promoters from other plant species and other organisms, such as insects, fungi, viruses, mammals, and the like, may also be suitable, and at times preferred. The promoter may be constitutive or inducible, or may be active in a certain tissue or tissues (tissue type-specific promoter), in a certain cell or cells (cell-type specific promoter), or at a particular stage or stages of development (development-type specific promoter). The choice of a promoter depends at least in part upon the application. Many promoters have been identified and isolated (*e.g.,* CaMV 35S promoter, maize ubiquitin promoter) (*see, generally*, GenBank and EMBL databases). Other promoters may be isolated by well-known methods. For example, a genomic clone for a particular gene can be isolated by probe hybridization. The coding region is mapped by restriction mapping, DNA sequence analysis, RNase probe protection, or other suitable method. The genomic region immediately upstream of the coding region comprises a promoter region and is isolated. Generally, the promoter region is located in the first 200 bases upstream, but may extend to 500 or more bases. The candidate region is inserted in a suitable vector in operative linkage with a reporter gene, such as in pBI121 in place of the CaMV 35S promoter, and the promoter is tested by assaying for the reporter gene after transformation into a plant cell. (*see, generally,* Ausubel *et al.*, *supra*; Sambrook *et al.*, *supra*; *Methods in Plant Molecular Biology and Biotechnology*, Ed. Glick and Thompson, CRC Press, 1993.)

[0095] Preferably, the vector contains a selectable marker for identifying transformants. The selectable marker preferably confers a growth advantage under appropriate conditions. Generally, selectable markers are drug resistance genes, such as neomycin phosphotransferase. Other drug resistance genes are known to those in the art and may be readily substituted. Selectable markers include ampicillin resistance, tetracycline resistance, kanamycin resistance, chloramphenicol resistance, and the like. The selectable marker also preferably has a linked constitutive or inducible promoter and a termination sequence, including a polyadenylation signal sequence. Other selection systems, such as positive selection can alternatively be used. Because the fungal *gus* genes of the present invention cleave Cba, they are particularly suitable for use as a positive selection marker.

[0096] The sequence of nucleotides encoding a β-glucuronidase may also include a classical secretion signal, whereby the resulting peptide is a precursor protein processed and secreted. Suitable signal sequences of plant genes include, but are not limited to the signal sequences from glycine-rich protein and extensin. In addition, a glucuronide permease gene to facilitate uptake of glucuronides may be co-transfected either from the same vector containing fungal GUS or from a separate expression vector.

[0097] A general vector suitable for use in the present invention is based on pCAMBIA 1305.2. Other vectors have been described (U.S. Patent Nos. 4,536,475; 5,733,744; 4,940,838; 5,464,763; 5,501,967; 5,731,179) or may be constructed based on the guidelines presented herein. The plasmid contains a left and right border sequence for integration into a plant host chromosome and also contains a bacterial origin of replication and selectable marker. These border sequences flank two genes. One is a kanamycin resistance gene (neomycin phosphotransferase) driven by a nopaline synthase promoter and using a nopaline synthase polyadenylation site. The second is the *E. coli gus* gene (reporter gene) under control of the CaMV 35S promoter and polyadenlyated using a nopaline synthase polyadenylation site. The *E. coli gus* gene is replaced with a gene encoding a fungal *gus* gene, especially one that cleaves Cba. If appropriate, the CaMV 35S promoter is replaced by a different promoter. Either one of

the expression units described above is additionally inserted or is inserted in place of the CaMV promoter and *gus* gene.

[0098] Plants may be transformed by any of several methods. For example, plasmid DNA may be introduced by *Agrobacterium* co-cultivation (*e.g.,* U.S. Patent No. 5,591,616; 4,940,838) or bombardment (*e.g.,* U.S. Patent No. 4,945,050; 5,036,006; 5,100,792; 5,371,015). Other transformation methods include electroporation (U.S. Patent No. 5,629,183), $CaPO_4$-mediated transfection, gene transfer to protoplasts (AU B 600221), microinjection, and the like (*see, Gene Transfer to Plants*, Ed. Potrykus and Spangenberg, Springer, 1995, for procedures). Preferably, vector DNA is first transfected into *Agrobacterium* and subsequently introduced into plant cells. Most preferably, the infection is achieved by *Agrobacterium* co-cultivation. In part, the choice of transformation methods depends upon the plant to be transformed. Tissues can alternatively be efficiently infected by *Agrobacterium* utilizing a projectile or bombardment method. Projectile methods are generally used for transforming sunflowers and soybean. Bombardment is often used when naked DNA, typically *Agrobacterium* binary plasmids or pUC-based plasmids, is used for transformation or transient expression.

[0099] Briefly, co-cultivation is performed by first transforming *Agrobacterium* by freeze-thaw method (Holsters *et al.*, *Mol. Gen. Genet. 163*: 181-187, 1978) or by other suitable methods (*see,* Ausubel, *et al. supra*; Sambrook *et al.*, *supra*). Briefly, a culture of *Agrobacterium* containing the plasmid is incubated with leaf disks, protoplasts, meristematic tissue, or calli to generate transformed plants (Bevan, *Nucl. Acids. Res. 12*:8711, 1984) (U.S. Patent No. 5,591,616). After co-cultivation for about 2 days, bacteria are removed by washing and plant cells are transferred to plates containing antibiotic (*e.g.,* cefotaxime) and a selective agent, such as Cba. Plant cells are further incubated for several days. The presence of the transgene may be tested for at this time. After further incubation for several weeks in selecting medium, calli or plant cells are transferred to regeneration medium and placed in the light. Shoots are transferred to rooting medium and then into glass house.

[0100] Briefly, for microprojectile bombardment, cotyledons are broken off to produce a clean fracture at the plane of the embryonic axis, which are placed broken surface up on medium with growth-regulating hormones, minerals and vitamin additives. Explants from other tissues or methods of preparation may alternatively be used. Explants are bombarded with gold or tungsten microprojectiles by a particle acceleration device and cultured for several days in a suspension of transformed *Agrobacterium.* Explants are transferred to medium lacking growth regulators but containing drug for selection and grown for 2-5 weeks. After 1-2 weeks more without drug selection, leaf samples from green, drug-resistant shoots are grafted to *in vitro* grown rootstock and transferred to soil. Classical tests for a transgene such as Southern blotting and hybridization or genetic segregation can also be performed.

[0101] A positive selection system, for example based on cellobiuronic acid in a culture medium lacking a carbon source is preferably used (*see,* U.S. Patent No: 6,268,493.

[0102] Activity of secreted GUS is conveniently assayed in whole plants or in selected tissues using a glucuronide substrate that is readily detected upon cleavage. Glucuronide substrates that are colorimetric are preferred. Field testing of plants may be performed by spraying a plant with the glucuronide substrate and observing color formation of the cleaved product.

Expression in other organisms

[0103] A variety of other organisms are suitable for use in the present invention. For example, various fungi, including yeasts, molds, and mushrooms, insects, especially vectors for diseases and pathogens, and other animals, such as cows, mice, goats, birds, aquatic animals (*e.g.,* shrimp, turtles, fish, lobster and other crustaceans), amphibians and reptiles and the like, may be transformed with a *gus* transgene.

[0104] The principles that guide vector construction for bacteria and plants, as discussed above, are applicable to vectors for these organisms. In general, vectors are well known and readily available. Briefly, the vector should have at least a promoter functional in the host in operative linkage with *gus*. Usually, the vector will also have

one or more selectable markers, an origin of replication, a polyadenylation signal and a transcription terminator.

[0105] The sequence of nucleotides encoding a β-glucuronidase may also include a classical secretion signal, whereby the resulting peptide is a precursor protein processed and secreted. Suitable secretion signals may be obtained from a variety of genes, such as mat-alpha or invertase genes. In addition, a permease gene may be co-transfected.

[0106] One of ordinary skill in the art will appreciate that a variety of techniques for producing transgenic animals exist. In this regard, the following U.S. patents teach such methodologies and are thus incorporated herein by reference: U.S. Patent Nos. 5,162,215; 5,545,808; 5,741,957; 4,873,191; 5,780,009; 4,736,866; 5,567,607; and 5,633,076.

Uses of fungal β-glucuronidase

[0107] As noted above, fungal β-glucuronidase may be used in a variety of applications. In certain aspects, fungal β-glucuronidase can be used as a reporter/effector molecule and as a diagnostic tool. As taught herein, fungal β-glucuronidase that cleaves Cba is preferred as an *in vivo* reporter/effector molecule, whereas, in *in vitro* diagnostic applications, the biochemical characteristics of the β-glucuronidase disclosed herein (*e.g.*, thermal stability, high turnover number) may provide preferred advantages.

[0108] Fungal GUS, either secreted or non-secreted, can be used as a marker/effector for transgenic constructions. In a certain embodiments, the transgenic host is a plant, such as rice, corn, wheat, or an aquatic animal. The transgenic GUS may be used in at least three ways: one in a method of positive selection, obviating the need for drug resistance selection, a second as a system to target molecules to specific cells, and a third as a means of detecting and tracking linked genes.

[0109] For positive selection, a host cell, (*e.g.*, plant cells) is transformed with a *gus* transgene (preferably coding for a secretable GUS). Selection is achieved by providing the cells with a glucuronidated form of a required nutrient (U.S. Patent Nos 5,994,629;

5,767,378; PCT US99/17804). For example, all cells require a carbon source, such as glucose. In one embodiment, glucose is provided as glucuronyl glucose (cellobiuronic acid), which is cleaved by GUS into glucose plus glucuronic acid. The glucose would then bind to transporters and be taken up by cells. The aglycone part of the glucuronide can be any required compound, including without limitation, a cytokinin, auxin, vitamin, carbohydrate, nitrogen-containing compound, and the like. It will be appreciated that this positive selection method can be used for cells and tissues derived from diverse organisms, such as animal cells, insect cells, fungi, and the like. The choice of glucuronide will depend in part upon the requirements of the host cell.

[0110] As a marker/effector molecule, secreted GUS (s-GUS) is preferred because it is non-destructive, that is, the host does not need to be destroyed in order to assay enzyme activity. A non-destructive marker has special utility as a tool in plant breeding. The GUS enzyme can be used to detect and track linked endogenous or exogenously introduced genes. GUS may also be used to generate sentinel plants that serve as bioindicators of environmental status. Plant pathogen invasion can be monitored if GUS is under control of a pathogen promoter. In addition, such transgenic plants may serve as a model system for screening inhibitors of pathogen invasion. In this system, GUS is expressed if a pathogen invades. In the presence of an effective inhibitor, GUS activity will not be detectable. In certain embodiments, GUS is co-transfected with a gene encoding a glucuronide permease.

[0111] Transgenes for introduction into plants encode proteins that affect fertility, including male sterility, female fecundity, and apomixis; plant protection genes, including proteins that confer resistance to diseases, bacteria, fungus, nematodes, herbicides, viruses and insects; genes and proteins that affect developmental processes or confer new phenotypes, such as genes that control meristem development, timing of flowering, cell division or senescence (*e.g.*, telomerase), toxicity (*e.g.*, diphtheria toxin, saporin), affect membrane permeability (*e.g.*, glucuronide permease (U.S. Patent No. 5,432,081)), transcriptional activators or repressors, alter nutritional quality, produce vaccines, and the like.

[0112] Insect and disease resistance genes are well known. Some of these genes are present in the genome of plants and have been genetically identified. Others of these genes have been found in bacteria and are used to confer resistance. Particularly well known insect resistance genes are the crystal genes of *Staphylococcus thuringiensis*. The crystal genes are active against various insects, such as lepidopterans, *Diptera*, *Hemiptera* and *Coleoptera*. Many of these genes have been cloned. For examples, see, GenBank; U.S. Patent Nos. 5,317,096; 5,254,799; 5,460,963; 5,308,760, 5,466,597, 5,2187,091, 5,382,429, 5,164,180, 5,206,166, 5,407,825, 4,918,066.

[0113] Other resistance genes to *Sclerotinia*, cyst nematodes, tobacco mosaic virus, flax and crown rust, rice blast, powdery mildew, verticillum wilt, potato beetle, aphids, as well as other infections, are useful within the context of this invention. Examples of such disease resistance genes may be isolated from teachings in the following references: isolation of rust disease resistance gene from flax plants (WO 95/29238); isolation of the gene encoding Rps2 protein from *Arabidopsis thaliana* that confers disease resistance to pathogens carrying the avrRpt2 avirulence gene (WO 95/28478); isolation of a gene encoding a lectin-like protein of kidney bean confers insect resistance (JP 71-32092); isolation of the Hm1 disease resistance gene to *C. carbonum* from maize (WO 95/07989); for examples of other resistance genes, see WO 95/05743; U.S. Patent No. 5,496,732; U.S. Patent No. 5,349,126; EP 616035; EP 392225; WO 94/18335; JP 43-20631; EP 502719; WO 90/11770; U.S. Patent 5,270,200; U.S. Patent Nos. 5,218,104 and 5,306,863). Nucleotide sequences for other transgenes, such as controlling male fertility, are found in U.S. Patent No. 5,478,369, references therein, and Mariani *et al.*, *Nature 347*:737, 1990.

[0114] In similar fashion, fungal GUS, can be used to generate transgenic insects for tracking insect populations or facilitate the development of a bioassay for compounds that affect molecules critical for insect development (*e.g.*, juvenile hormone). Secreted GUS may also serve as a marker for beneficial fungi destined for release into the environment. The non-destructive marker is useful for detecting persistence and competitive advantage of the released organisms.

[0115] In animal systems, secreted GUS may be used to achieve extracellular cleavage of glucuronides (*e.g,* pharmaceutical glucuronide) and examine conjugation patterns of glucuronides. Furthermore, as discussed above, secreted GUS may be used as a transgenic marker to track cells or as a positive selection system, or to assist in development of new bioactive GUS substrates that do not need to be transported across membrane. Aquatic animals are also suitable hosts for GUS transgene. GUS may be used in these animals as a marker or effector molecule.

[0116] Within the context of this invention, GUS may also be used in a system to target molecules to cells. This system is particularly useful when the molecules are hydrophobic and thus, not readily delivered. These molecules can be useful as effectors (*e.g.,* inducers) of responsive promoters. For example, molecules such as ecdysone are hydrophobic and not readily transported through phloem in plants. When ecdysone is glucuronidated it becomes amphipathic and can be delivered to cells by way of phloem. Targeting of compounds such as ecdysone-glucuronic acid to cells is accomplished by causing cells to express receptor for ecdysone. As ecdysone receptor is naturally only expressed in insect cells, however a host cell that is transgenic for ecdysone receptor will express it. The glucuronide containing ecdysone then binds only to cells expressing the receptor. If these cells also express GUS, ecdysone will be released from the glucuronide and able to induce expression from an ecdysone-responsive promoter. Plasmids containing ecdysone receptor genes and ecdysone responsive promoter can be obtained from Invitrogen (Carlsbad, CA). Other ligand-receptors suitable for use in this system include glucocorticoids/glucocorticoid receptor, estrogen/estrogen receptor, antibody and antigen, and the like (*see* also U.S. Patent Nos. 5,693,769 and 5,612,317).

[0117] In another aspect, purified fungal β-glucuronidase is used in medical applications. For these applications, secretion is not a necessary characteristic although it may be a desirable characteristic for production and purification. The biochemical attributes, such as the increased stability and enzymatic activity disclosed herein are preferred characteristics. The fungal glucuronidase preferably has one or more of the disclosed characteristics.

[0118] For the majority of drug or pharmaceutical analysis, the compounds in urine, blood, saliva, or other bodily fluids are de-glucuronidated prior to analysis. Such a procedure is undertaken because compounds are often, if not nearly always, detoxified by glucuronidation in vertebrates. Thus, drugs that are in circulation and have passed through a site of glucuronidation (*e.g.*, liver) are found conjugated to glucuronic acid. Such glucuronides yield a complex pattern upon analysis by, for example, HPLC. However, after the aglycone (drug) is cleaved from the glucuronic acid, a spectrum can be compared to a reference spectrum. Currently, *E. coli* GUS is utilized in medical diagnostics, but as shown herein, fungal GUS may have superior qualities.

[0119] The fungal GUS enzymes disclosed herein may be used in traditional medical diagnostic assays, such as described above for drug testing, pharmacokinetic studies, bioavailability studies, diagnosis of diseases and syndromes, following progression of disease or its response to therapy and the like (*see* U.S. Patent Nos. 5,854,009, 4,450,239, 4,274,832, 4,473,640, 5,726,031, 4,939,264, 4,115,064, 4,892,833). These β -glucuronidase enzymes may be used in place of other traditional enzymes (*e.g.*, alkaline phosphatase, horseradish peroxidase, β-galactosidase, and the like) and compounds (*e.g.*, green fluorescent protein, radionuclides) that serve as visualizing agents. Fungal GUS has qualities advantageous for use as a visualizing agent: it is highly specific for the substrate, water soluble and the substrates are stable. Thus, fungal GUS is suitable for use in Southern analysis of DNA, Northern analysis, ELISA, and the like.

[0120] In preferred embodiments, fungal GUS binds a hapten, either as a fusion protein with a partner protein that binds the hapten (*e.g.*, avidin that binds biotin, antibody) or alone. If used alone, fungal GUS can be mutagenized and selected for hapten-binding abilities. Mutagenesis and binding assays are well known in the art. In addition, fungal GUS can be conjugated to avidin, streptavidin, antibody or another hapten-binding protein and used as a reporter in the myriad of assays that currently employ enzyme-linked binding proteins. Such assays include immunoassays, Western blots, *in situ* hybridizations, HPLC, high-throughput binding assays, and the like (see, for examples, U.S. Patent Nos. 5,328,985 and 4,839,293, which teach avidin and

streptavidin fusion proteins and U.S. Patent No. 4,298,685, Diamandis and Christopoulos, *Clin. Chem. 37*:625, 1991; Richards, *Methods Enzymol. 184*:3, 1990; Wilchek and Bayer, *Methods Enzymol. 184*:467, 1990; Wilchek and Bayer, *Methods Enzymol. 184*:5, 1990; Wilchek and Bayer, *Methods Enzymol. 184*:14, 1990; Dunn, *Methods Mol. Biol. 32*:227, 1994; Bloch, *J. Hitochem. Cytochem. 41*:1751, 1993; Bayer and Wilchek *J. Chromatogr. 510*:3, 1990, which teach various applications of enzyme-linked technologies and methods).

[0121] Fungal GUSes can also be used in therapeutic methods. By turning compounds such as drugs into glucuronides, the compound is inactivated. When a glucuronidase is expressed or targeted to the site for delivery, the glucuronide is cleaved and the compound delivered. For these purposes, GUS may be expressed as a transgene or delivered, for example, coupled to an antibody specific for the target cell (*see e.g.*, U.S. Patent Nos. 5,075,340, 4,584,368, 4,481,195, 4,478,936, 5,760,008, 5,639,737, 4,588,686).

[0122] The present invention also provides kits comprising fungal GUS protein or expression vectors containing fungal *gus* gene. One exemplary type of kit is a dipstick test. Such tests are widely utilized for establishing pregnancy, as well as other conditions. Generally, these dipstick tests assay the glucuronide form, but it would be advantageous to use reagents that detect the aglycone form. Thus, GUS may be immobilized on the dipstick adjacent to or mixed in with the detector molecule (*e.g.*, antibody). The dipstick is then dipped in the test fluid (*e.g.*, urine) and as the compounds flow past GUS, they are cleaved into aglycone and glucuronic acid. The aglycone is then detected. Such a setup may be extremely useful for testing compounds that are not readily detectable as glucuronides.

[0123] In a variation of this method, the fungal GUS enzyme is engineered to bind a glucuronide, but lacks enzymatic activity. The enzyme will then bind the glucuronide and the enzyme is detected by standard methodology. Alternatively, GUS is fused to a second protein, either as a fusion protein or as a chemical conjugate that binds an aglycone. The fusion is incubated with the test substance and an indicator substrate is

added. This procedure may be used for ELISA, Northern, Southern analysis and the like.

[0124] The following examples are offered by way of illustration, and not by way of limitation.

EXAMPLES

EXAMPLE 1

IDENTIFICATION OF FUNGI EXPRESSING β-GLUCURONIDASE

[0125] In this example, fungi are screened for expression of β-glucuronidase by a colorimetric assay. Blue-staining fungi are selected, purified, and identified by comparison of rRNA sequences to known sequences.

[0126] Soil samples from around Canberra, Australia, are shaken for 15 sec in 500 μL of sterile water. After centrifugation at 17,000 × g for 15 s, 100 μL of the supernatant are plated on modified M9 medium containing 4-O-(β-D-glucuronyl)-D-glucose (cellobiouronic acid; Cba) as the sole carbon source and 5-bromo-4-chloro-3-indolyl-β-D-glucuronide (X-GlcA) as an indicator substrate for β-glucuronidases (1.28 % $Na_2HPO_4 \cdot 7\ H_2O$, 0.3 % KH_2PO_4, 0.05 % NaCl, 0.1 % NH_4Cl, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 20 $\mu g\ L^{-1}$ folic acid, 20 $\mu g\ L^{-1}$ biotin, 50 $\mu g\ L^{-1}$ nicotinic acid, 50 $\mu g\ L^{-1}$ riboflavin, 50 $\mu g\ L^{-1}$ thiamin·Cl, 0.1 mg L^{-1} pyridoxine·Cl, 2.8 mg L^{-1} H_3BO_3, 1.8 mg L^{-1} $MnCl_2 \cdot 4\ H_2O$, 1.4 mg L^{-1} $FeSO_4 \cdot 7\ H_2O$, 30 $\mu g\ L^{-1}$ $CuCl_2 \cdot 2\ H_2O$, 20 $\mu g\ L^{-1}$ $CoCl_2 \cdot 6\ H_2O$, 3 mg L^{-1} Na_2-EDTA, 10 mM Cba, 50 mg L^{-1} X-GlcA, 1.5 % agar). Cba is used to enrich for microorganisms with β-glucuronidase activity, because they should be able to hydrolyze Cba and thus grow on it as the sole carbon source. Such microorganisms are expected to stain blue as a result of hydrolyzing X-GlcA.

[0127] Blue-staining fungi are selected for further analysis. They are purified from any bacteria that may be adhering to fungi by consecutive sub-cultivations on YPD plates containing a combination of anti-bacterial antibiotics (1% yeast extract, 2% peptone, 2% glucose, 50 mg L^{-1} ampicillin, 50 mg L^{-1} streptomycin, 50 mg L^{-1} nalidixic acid). After a minimum of six rounds of sub-cultivations, the isolates are transferred back on the original Cba medium to confirm their β-glucuronidase activity.

[0128] Purified isolates are grown in liquid YPD medium at 29°C. Genomic DNA is isolated from hyphae using the DNAzol kit (Invitrogen; Carlsbad, CA, USA). A region between the 18S and 26S rRNA genes is amplified with primers ITS-fwd1 (SEQ

ID NO:24) and ITS-rev4 (SEQ ID NO:25) (Table 1). This region contains the 5.8S rRNA gene flanked by intergenic transcribed spacers 1 and 2 (ITS1, ITS2), which are highly variable and well suited for identification of fungal isolates at the species level. In case no perfect match with a known species is found, a region of the 18S-rRNA gene is amplified using primers NS3 (SEQ ID NO:26) and NS6 (SEQ ID NO:27) for a tentative phylogenetic placement at the genus level (Table 1).

Table 1. Primers for amplification of fungal ITS regions and 18S rRNA gene fragments.

Primer	No. bases	T_m (°C)	Sequence	SEQ ID NO:
ITS-fwd1	19	56	5'-TCCGTAGGTGAACCTGCGG-3'	24
ITS-rev4	20	50	5'-TCCTCCGCTTATTGATATGC-3'	25
NS3	21	62	5'-GCAAGTCTGGTGCCAGCAGCC-3'	26
NS6	24	57	5'-GCATCACAGACCTGTTATTGCCTC-3'	27

[0129] The ITS region is amplified from 50 ng of genomic DNA with 0.5 U of REDTaq (Sigma; St. Louis, Missouri, USA) in 20 μL of 10 mM Tris (pH 8.3), 50 mM KCl, 2.5 mM $MgCl_2$, 0.25 μM dNTPs, 1 μM ITS-fwd1, and 1 μM ITS-fwd4. After initial denaturation at 94°C for 2 min, the reactions are cycled 35 times at 94°C (20 sec), 50°C (40 sec), and 72°C (1.5 min), followed by a final extension at 72°C for 5 min. The 18S rRNA gene fragment is amplified under identical conditions with NS3 and NS6 primers and an annealing temperature of 60°C. Amplified fragments are separated on a 1.2% TAE agarose gel, excised, and extracted using a gel nebulizer (Ultrafee-DA; Millipore; Bedford, Maryland, USA). Two microliters of each of the extracted amplified products are then sequenced using the BigDye Terminator sequencing mix (Perkin Elmer ABI, Poster City, CA, USA). Cycling conditions for the sequence reactions are: 25 cycles of 96°C (30 sec), 50°C (15 sec), and 60°C (4 min). After precipitation of the cycling products with 4 volumes of 75 % isopropanol, they are separated on a polyacrylamide gel to obtain their nucleotide sequences.

Table 2. Sequences of rRNA genes

ITS1 – 5.8S rRNA gene – ITS2 – 28S rRNA gene (partial) of *Penicillium canescens* isolate RPK (SEQ ID NO:28)

CGAGAATTCTCTGAATTCAACCTCCCACCCGTGTTTATTGTACCTTGTTGCTTCGGCGGGCCCGCCTCAC
GGCCGCCGGGGGGCATCTGCCCCCGGGCCCGCGCCCGCCGAAGACACCTTGAACTCTGTATGAAAATTGC
AGTCTGAGTCTAAATATAAATTATTTAAAACTTTCAACAACGGATCTCTTGGTTCCGGCATCGATGAAGA
ACGCAGCGAAATGCGATACGTAATGTGAATTGCAGAATTCAGTGAATCATCGAGTCTTTGAACGCACATT
GCGCCCCCTGGTATTCCGGGGGGCATGCCTGTCCGAGCGTCATTGCTGCCCTCAAGCCCGGCTTGTGTGT
TGGGTCTCGTCCCCCTTCCCGGGGGGACGGGCCCGAAAGGCAGCGGCGGCACCGCGTCCGGTCCTCGAGC
GTATGGGGCTTTGTCACCCGCTCTGTAGGCCCGGCCGGCGCTTGCCGATCAACCAAAACTTTTTTCCAGG
TTGACCTCGGATCAGGTAGGGATACCCGCTGAACTTAA

ITS1 – 5.8S rRNA gene – ITS2 – 28S rRNA gene (partial) of *Scopulariopsis* sp. isolate RP38.3 (SEQ ID NO:29)

GGGATCATTACCGAAGTTACTCTTCAAAACCCATTGTGAACCTTACCTCTTGCCGCGCGTTGCCTCGGCG
GGGAGGCGGGGTCTGGGTCGGCGCGCCCCTCACCGGGCCGCCGTCCCGTCCCGTCCCCGCCGGCCGCGCC
AAACTCTAAATTTGAAAAAGCGTACTGCACGTTCTGATTCAAAACAAAAAACAAGTCAAAACTTTTAACA
ACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGCGATAAGTAATGTGAATTGCAGAATT
CAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCGGCAGCAATCTGCCGGGCATGCCTGTCCGAGCG
TCATTTCTTCCCTCGAGCGCGGCTAGCCCTACGGGGCCTGCCGTCGCCCGGTGTTGGGGCTCTACGGGTG
GGGCTCGTCCCCCCCGCAGTCCCCGAAATGTAGTGGCGGTCCAGCCGCGGCGCCCCCTGCGTAGTAGATC
CTACATCTCGCATCGGGTCCCGGCGAAGGCCAGCCGTCGAACCTTTTATTTCATGGTTTGACCTCGGATC
AGGTAGGGTTACCCGCT

18S rRNA gene (partial) of *Penicillium canescens* isolate RPK (SEQ ID NO:30)

TTCCAGCTCCAATAGCGTATATTAAAGTTGTTGCAGTTAAAAAGCTCGTAGTTGAACCTTGGGTCTGGCT
GGCCGGTCCGCCTCACCGCGAGTACTGGTCCGGCTGGACCTTTCCTTCTGGGGAACCTCATGGCCTTCAC
TGGCTGTGGGGGGAACCAGGACTTTTACTGTGAAAAAATTAGAGTGTTCAAAGCAGGCCTTTGCTCGAAT
ACATTAGCATGGAATAATAGAATAGGACGTGCGGTTCTATTTTGTTGGTTTCTAGGACCGCCGTAATGAT
TAATAGGGATAGTCGGGGGCGTCAGTATTCAGCTGTCAGAGGTGAAATTCTTGGATTTGCTGAAGACTAA
CTACTGCGAAAGCATTCGCCAAGGATGTTTTCATTAATCAGGGAACGAAAGTTAGGGGATCGAAGACGAT
CAGATACCGTCGTAGTCTTAACCATAAACTATGCCGACTAGGGATCGGACGGGATTCTATAATGACCCGT
TCGGCACCTTACGAGAAATCAAAGTTTTTGGGTTCTGGGGGGAGTATGGTCGCAAGGCTGAAACTTAAAG
AAATTGACGGAAGGGCACCACAAGGCGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGGAAACTCACC
AGGTCCAGACAAAATAAGGATTGACAGATTGAGAGCTCTTTCTTGATCTTTTGGATGGTGGTGCATGGCC
GTTCTTAGTTGGTGGAGTGATTTGTCTGCTTAATTGCGATAACGAACGAGACCTCGGCCCTTAAATAGCC
CGGTCCGCATTTGCGGGCCGCTGGCTTCTTAGGGGGACTATCGGCTCAAGCCGATGGAAGTGCAGG

18S rRNA gene (partial) of *Scopulariopsis* sp. isolate RP38.3 (SEQ ID NO:31)

AATTCCAGCTCCAATAGCGTATATTAAAGTTGTTGTGGTTAAAAAGCTCGTAGTCGAACCTTGGGCCTGG
CTGGCCGGTCCCCCTCACCGGGTGCACTGATCCAGCCGGGCCTTTCCCTCTGTGGAACCCCATGGCCTTC
ACTGGCTGTGCGGGGGAAACAGGACTTTTACTGTGAAAAAATTAGAGTGCTCCAGGCAGGCCTATGCTCG
AATACATTAGCATGGAATAATAGAATAGGACGTGTGGTTCTATTTTGTTGGTTTCTAGGACCGCCGTAAT
GATTAATAGGGACAGTCGGGGGCATCAGTATTCAGTTGTCAGAGGTGAAATTCTTGGATCTACTGAAGAC
TAACTACTGCGAAAGCATTTGCCAAGGATGTTTTCATTGATAAGGAACGAAAGTTAGGGGATCGAAGACG
ATCAGATACCGTCGTAGTCTTAACTATAAACTATGCCGACTAGGGATCGGACGATGTTATTATTTGACGC
GTTCGGCACCTTTCGAGAAATCAAAGTGCTTGGGCTCCAGGGGGAGTATGGTCGCAAGGCTGAAACTTAA
AGAAATTGACGGAAGGGCACCACCAGGGGTGGAACCTGCGGCTTAATTTGACTCAACACGGGGAAACTCA
CCAGGTCCAGACACAGTGAGGATTGACAGATTGAGAGCTCTTTCTTGATTCTGTGGGTGGTGGTGCATGG
CCGTTCTTAGTTGGTGGAGTGATTTGTCTGCTTAATTGCGATAACGAACGAGACCTTAACCTGCTAAATA
GCCCGTACTGCTCTGGCAGTTCGCCGGCTTCTTAGAGGGACTATCGGCTCAAGCCGAGGAAT

[0130] The ITS sequences (isolate RP38.3, (SEQ ID NO:28); isolate RPK (SEQ ID NO:29)) are then subjected to a similarity search, using the BLAST 2.0 server at NCBI (http://www.ncbi.nlm.nih.gov:80/BLAST; Altschul et al., *J Mol Biol* 215: 403-410, 1990). The sequences of the 18S rRNA genes (isolate RP 38.3, (SEQ ID NO:30); isolate RPK (SEQ ID NO:31)) are aligned against other eukaryotic 18S rRNA genes, using the facilities of the Ribosomal Database Project at Michigan State University (http://rdp.cme.msu.edu/html/; Maidak et al., *Nucleic Acids Res* **28**: 173-174, 2000). The deduced phylogenetic placement of isolated fungi with *β*-glucuronidase activity is shown in Table 3.

Table 3

Phylogenetic placement of fungal and bacterial isolates with β-glucuronidase activity.

Type of Isolate	ID	ITS1 – 5.8S rRNA gene – ITS2 – 28S rRNA gene (partial)		SSU rRNA gene (partial)	
		Closest match	Homology	Closest match	Homology
Fungus	RP38.3	_ [a]	–	*Scopulariopsis brevicaulis* (AY083220)	99.8% (826/828)
Fungus	RPK	*Penicillium canescens* (AF033493)	100% (528/528)	*Penicillium sacculum* [b] (AB027410)	99.9% (832/833)
				Penicillium herquei (AB086834)	
				Eladia saccula (AB031391)	
				Eupenicillium sp. (AY297772)	

[a] No continuous match spanning both ITS and the 5.8S rRNA gene.
[b] The database does not contain the *Penicillium canescens* sequence.

[0131] Based on the results shown in Table 2, it is concluded that the two characterized GUS-expressing fungi belong to the Pezizomycotina (= Euascomycetes) subphylum of the Ascomycota phylum of the fungi kingdom. One of them (*Penicillium canescens*) is member of the Eurotiomycetes class, while the other (*Scopulariopsis breviaulis*) is member of the Sordariomycetes class. In addition, a second isolate of *Penicillium canescens*, DSM 1215, that expresses β-glucuronidase is identified by similar methods.

EXAMPLE 2

BIOCHEMICAL CONFIRMATION OF β-GLUCURONIDASE ACTIVITY IN FUNGI

[0132] In this example, enzyme activity of β-glucuronidase in fungi is quantified following growth in media containing different inducers or no inducer of expression.

[0133] GUS-expressing fungi are isolated based on their ability to hydrolyze X-GlcA, a widely used GUS substrate. To confirm β-glucuronidase activity, the hydrolysis of 4-methylumbelliferyl-β-D-glucuronide (MU-GlcA), another widely used GUS substrate, is measured *in vitro*. Both purified fungal isolates are grown in liquid YPD medium on a shaker at 200 rpm/29°C for 3 days. Hyphal aggregates are then vacuum-filtered, washed once with modified M9 medium lacking Cba (see Example 1), and suspended in the same medium. After 6 h of starvation in this medium, putative inducers of β-glucuronidase activity are added. These include X-GlcA (0.1 mM), Cba (20 mM) and glucuronic acid (GlcA; 0.1 and 20 mM). The fungi are then incubated in these media for an additional 6 h, in the course of which aliquots of hyphal aggregates are taken, vacuum-filtered and ground in liquid nitrogen.

[0134] Proteins are extracted in 40 mM PIPES pH 7.0, 2 mM di-thiothreitol, 1 mM ethylenediaminetetraacetic acid, 1 mM phenylmethylsulfonyl-fluoride, 0.1 % [v/v] Triton X-100. Protein concentrations in the supernatants obtained by centrifugation at 23,000×g for 15 min/4°C are determined with the Bradford assay, using bovine serum albumin dissolved in extraction buffer as a standard. The β-glucuronidase activity of these extracts is then measured in 160 μL of extraction buffer to which 0.1 mg mL^{-1} BSA, 0.1 % Triton X-100, 1 mM MU-GlcA and 3 μg mL^{-1} of extracted proteins had been added. The reactions are incubated at 30°C for increasing periods of time and stopped by addition of 40 μL of 2 M Na_2CO_3. The amount of 4-methylumelliferone (MU) released from MU-GlcA is quantified fluorimetrically with a SpectraFluorPlus microplate reader (excitation: 360 nm, emission: 465 nm; Tecan GmbH; Grödig, Austria), using MU dissolved in assay buffer as a standard.

[0135] Figure 1 shows that GUS activity is only detectable if glucuronides such as X-GlcA and Cba, or free glucuronic acid, are added to the growth medium. After their addition, the GUS activity increases in a time-dependent manner. In the case where no

inducer is added, the GUS activity remains below the detection limit. These data confirm that the isolated fungi express the enzyme GUS and hydrolyze glucuronides.

EXAMPLE 3

CLONING OF FUNGAL *GUS* GENES

[0136] Isolated genomic DNA from three fungal isolates is used as a template to amplify fragments of gus genes using degenerate primers. These primers are designed based on a multiple alignment of known gus genes from bacteria and animals. They are predicted to amplify a 1.2 kb-long fragment of an intron-less gene. The sequences of the primers are given in Table 4. PCR amplification is carried out in 20 μL of 10 mM Tris (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.25 μM dNTPs, 1 μM gus-fwd+T3, 1 μM gus-rev+T7, containing 0.5 U of REDTaq (Sigma; St. Louis, Missouri, USA) and 50 ng of genomic DNA. Cycling conditions are 94°C (2 min), followed by 35 cycles of 94°C (20 sec), 48°C (40 sec) and 72°C (2 min 30 sec), and a final extension at 72°C for 7 min.

Table 4. Degenerate primers used to amplify a 1.2 kb fragment of fungal gus genes.*

Primer	No. of bases	Sequence
gus-fwd+T3	39	5'-AATTAACCCTCACTAAAGGGAYTTYTWYAAYTAYGCIGG (SEQ ID NO:32)
gus-rev+T7	39	5'-GTAATACGACTCACTATAGGGRAARTCIGCRAARAACCA (SEQ ID NO:33)

*T3 and T7 handles are underlined

[0137] A distinct 1.2 kb band is obtained from all three GUS-expressing fungal isolates, suggesting suggests that none of the *gus* fragments contains an intron. The bands are extracted from the gel and sequenced with T3 and T7 primers as described in Example 1. Hypothetical protein sequences, generated by translation of the obtained sequences in all three reading frames, are subjected to a similarity search as described in Example 1 to confirm that the amplified DNA fragments are derived from gus genes.

[0138] The complete nucleotide sequences of the three gus genes, part of their promoters, and the downstream regions are obtained by Semi-Random Two-Step PCR (Chun et al., *Yeast* 13: 233-240, 1997). The sequences are presented in Figures 2-4 (SEQ ID NOs:1, 3, and 5). In addition, limited upstream and downstream sequence was obtained. There is a small (20 bp) insertion in the downstream sequence of the DSM isolate compared to the RPK isolate. The remainder of those two sequences are about 90% identical.

EXAMPLE 4

SEQUENCE ANALYSES OF FUNGAL *GUS* GENES AND THEIR PRODUCTS

[0139] Figures 2A-C (SEQ ID NO:1) and 3A-C (SEQ ID NO:3) display the nucleotide sequences of the genomic fragments isolated from two of the three fungal isolates with β-glucuronidase activity. Each fragment contains one continuous open reading frame (ORF). There are no ribosomal binding sites (GAGGA) situated 8 to 13 nucleotides upstream of the initiation codon of bacterial genes. Instead, several features of eukaryotic genes are present. In each case the predicted transcriptional start (see arrows in Figures 2 and 3), situated at an adenine nucleotide, is surrounded by pyrimidine nucleotides (one upstream and four to five downstream; Figures 2 and 3). This is a typical feature of eukaryotic genes (Knippers et al., Molekulare Genetik (5. Auflage). Georg Thieme Verlag, Stuttgart, Germany, 1990). TATA box-like motifs are located at -40 bp (*Scopulariopsis* sp.) or -32 bp and -19 bp (*Penicillium canescens*). In *Scopulariopsis* sp., the TATA box-like motif is surrounded by guanine nucleotides in positions characteristic for eukaryotic genes (Knippers et al., *supra*). A Kozak sequence-like motif (CCACC), known to enhance translation, is located immediately upstream of the initiation codon of the *Scopulariopsis* sp. gene (Kozak, *Cell* **44**: 283-292, 1986). The 3' untranslated region of the gus genes contains putative polyadenylation signals and sites, which in the case of *Scopulariopsis* sp. exhibits a perfect match with the consensus sequences described in the literature (AATAAA with CA at +12 bp; Watson et al., Molecular Biology of the Gene (4^{th} edition). The Benjamin/Cummings Publishing Company, Inc., CA, USA 1987). In addition, the

promoter region of gus in *Scopulariopsis* sp. contains three poly(dA) stretches characteristic for housekeeping genes in *Saccharomyces cerevisiae*, which is also a member of the Ascomycota phylum (Watson et al., *supra*).

[0140] The ORFs are translated into amino acid sequences (see Figure 2, SEQ ID NO:2 and Figure 3 SEQ ID NO:4). Analysis using a neural-network program trained to recognize eukaryotic N-terminal signal peptides, reveals that both fungal GUS proteins contain signal peptides (SignalP V1.1 at http://www.cbs.dtu.dk/services/SignalP; Nielsen et al., *Protein Engineering* **12**: 3-9, 1999). The predicted cleavage positions are between amino acids No. 26 and 27 (*Scopulariopsis* sp.) or 18 and 19 (*Penicillium canescens*). The presence of these N-terminal signal peptides suggests that both fungal isolates may produce secreted β-glucuronidases. This is consistent with the observation that both stain the surrounding agar blue.

[0141] The protein sequences are subjected to a similarity search, using the BLASTP program at the BLAST 2.0 server of NCBI (http://www.ncbi.nlm.nih.gov:80/BLAST; Altschul et al., *J Mol Biol* 215: 403-410, 1990). Results of these analyses demonstrate that the gene products are closely related to fungal and mammalian β-glucuronidases (e values range from 10^{-180} to 10^{-53}). A conserved domain (CD) search at the same server identifies three CDs: pfam02837 (glycosyl hydrolases family 2; sugar-binding domain), pfam02836 (glycosyl hydrolases family 2; TIM barrel domain), and pfam00703 (glycosyl hydrolases family 2; immunoglobulin-like β-sandwich domain). In addition, both fungal GUS proteins contain the two signatures that, according to the Swiss Institute of Bioinformatics, characterize family 2 glycosyl hydrolases (http://www.expasy.ch/cgi-bin/nicedoc.pl?PDOC00531). This confirms that fungal GUS proteins, like GUS proteins from other organisms, are members of family 2 of glycosyl hydrolases.

EXAMPLE 5

IDENTIFICATION OF ADDITIONAL FUNGAL *GUS* GENES THROUGH SEQUENCE MINING

[0142] To compare the amino acid sequences of the fungal GUS proteins with those of other β-glucuronidases, the sequences of other GUS proteins are retrieved from

GenBank (http://www.ncbi.nlm.nih.gov/Genbank/). In addition, using the TBLASTN program at the BLAST 2.0 server at NCBI (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?), fungal genomes are mined for non-annotated gus genes and translated into proteins.

[0143] In addition, the amino acid sequence of GUS of isolate RPK (*Penicillium canescens*) is used as query sequence to search for additional fungal gus genes in the Whole-Genome-Shotgun Sequences (WGS) database at NCBI. A TBLASTN search, carried out on 12 July 2003 (request ID 1057998419-03767-31842) identifies two more fungal genes in the genomes of *Aspergillus nidulans* and *Gibberella zeae* (anamorph: *Fusarium graminearum*). The gus gene of *Aspergillus nidulans* is located between positions 285949 and 287784 (frame +1) in the sequence deposited under GenBank accession number AACD01000093.1. The gus gene of *Gibberella zeae* is located between positions 77805 and 76006 (frame -3) in the sequence deposited under GenBank accession number AACM01000315.1. The DNA sequences (*G. zeae* (SEQ ID NO:7); *A. nidulans* (SEQ ID NO:9)) and predicted amino acid sequences (*G. zeae* (SEQ ID NO:8); *A. nidulans* (SEQ ID NO:10)) are presented in Figures 5 and 6. Similar to the gus genes of *Penicillium canescens* and *Scopulariopsis*, there are no introns in these genes. Both of these fungi belong to the Pezizomycotina (= Euascomycetes) subphylum of the Ascomycota phylum of the fungi kingdom. One of them (*Aspergillus nidulans*) is member of the Eurotiomycetes class, while the other (*Gibberella zeae*) is member of the Sordariomycetes class.

[0144] The predicted amino acid sequences of these two additional gus genes are used as query sequences in a similarity search using the BLASTP program at the BLAST 2.0 server of NCBI (http://www.ncbi.nlm.nih.gov:80/BLAST; Altschul et al., *J Mol Biol* 215: 403-410, 1990). Results of these analyses demonstrate that their gene products are closely related to fungal and mammalian β-glucuronidases (e values range from 10^{-174} to 10^{-79}). This search also identifies three CDs: pfam02837 (glycosyl hydrolases family 2; sugar-binding domain), pfam02836 (glycosyl hydrolases family 2; TIM barrel domain), and pfam00703 (glycosyl hydrolases family 2; immunoglobulin-like β-sandwich domain). Furthermore, both fungal GUS proteins contain the two

signatures that, according to the Swiss Institute of Bioinformatics, characterize family 2 glycosyl hydrolases (http://www.expasy.ch/cgi-bin/nicedoc.pl?PDOC00531).

[0145] The sequences of all GUS proteins are aligned with AlignX sofware (InforMax, Bethesda, MD, USA), which is based on the ClustalW program (Thompson et al., *Nucleic Acids Res* **22**: 4673-4680, 1994). BLOSUM 62 is chosen as the protein weight matrix (Henikoff and Henikoff, *Proc Natl Acad Sci USA* **89**: 10915-10919, 1992). The gap-opening penalty is adjusted to 10, the gap-extension penalty to 0.05, and the gap-separation distance to 8. An end gap-separation penalty and residue-specific and hydrophilic gap penalties are included. The resulting multiple alignment is displayed in Figure 7. This alignment shows considerable levels of sequence identity and similarity, particularly in the regions of the family 2 glycosyl hydrolase signatures. Two glutamate residues, at amino acids 562 and 607 as counted in the consensus sequence, which are previously shown to be required for catalytic activity of family 2 glycosyl hydrolases (Wong et al., *J Biol Chem* **273**: 34057-34062, 1998; Islam et al., *J Biol Chem* **274**: 23451-23455, 1999), are conserved in all GUS proteins, including the fungal forms (see asterisks in Figure 7).

[0146] In pair-wise alignments, the overall identity (similarity) to GUS^{Ecoli} is 49.6 % (60.5 %) for *Scopulariopsis* sp. and 50.3 % (61.6 %) for *Penicillium canescens*. The identities at the DNA level are 55.3 % (*Scopulariopsis* sp.) and 50.8 % (*Penicillium canescens*). The overall identity (similarity) to GUS^{Ecoli} is 47.3 % (59.1 %) for *Aspergillus nidulans* and 50.4 % (63.3 %) for *Gibberella zeae*. Like the *Penicillium* and *Scopulariopsis* GUS proteins, the gene product from *Aspergillus nidulans* has an N-terminal signal peptide with a predicted cleavage position between amino acid No. 20 and 21 (SignalP V1.1 at http://www.cbs.dtu.dk/services/SignalP; Nielsen et al., *Protein Engineering* **12**: 3-9, 1999). By contrast, the predicted gene product of *Gibberella zeae* does not appear to have an N-terminal signal peptide (Figure 7).

EXAMPLE 6

EXPRESSION OF FUNGAL *GUS* GENES IN *ESCHERICHIA COLI* AND RICE

[0147] To confirm that the isolated fungal gus genes indeed confer β-glucuronidase activity to organisms lacking it, the genes are cloned and transformed into a *gus*-deleted bacterium and a plant. The coding region of *gus* downstream of the predicted signal peptide cleavage site is amplified from genomic DNA of both *Penicillium canescens* and *Scopulariopsis* sp. Both pairs of forward and reverse primers contain restriction enzyme sites to facilitate subsequent cloning steps. Genomic DNA (5 – 50 ng) is used as a template in 20-μL amplification reactions containing 60 mM Tris·SO_4 (pH 9.1), 18 mM $(NH_4)_2SO_4$, 1.8 mM $MgSO_4$, 0.2 mM dNTPs, 0.2 μM fwd and reverse primers (Table 4) and 1 U of ELONGase (Invitrogen; Carlsbad, CA, USA). Cycling conditions are 94°C (30 sec), followed by 30 cycles of 94°C (20 sec) and 68°C (4 min), and a final extension at 68°C for 7 min. Amplified products are purified with the Qiagen PCR purification kit of (Qiagen GmbH; Hilden, Germany) and partially digested with SpeI and PmlI restriction enzymes. The digested fragments are separated on a TAE agarose gel (1.2 %) and extracted from the gel using the Qiagen Gel Extraction Kit.

Table 5

Primer	No. of bases	Sequence*
gus^{Scop}-fwd+SpeI	36	5'-CATAGCACTAGTGCCGACACTGACCAATGGAAGACG-3' (SEQ ID NO:34)
gus^{Scop}-rev+PmlI	35	5'-CGGTTACACGTGAGCACCGGAAGTACCGTTCCCCA-3' (SEQ ID NO:35)
gus^{Pcan}-fwd+SpeI	35	5'-CATAGCACTAGTACACCTGCAGCTCGGCACTTTCC-3' (SEQ ID NO:36)
gus^{Pcan}-rev+PmlI	64	5'-CGGTTACACGTGATTCTTATCAATACTAGTCCACCTTGCCCTCAAA-3' (SEQ ID NO:37)

* Restriction enzyme sites are underlined.

gus^{Scop} *Scopulariopsis gus*

gus^{Pcan} *Penicillium gus*

[0148] Aliquots are then ligated to a SpeI/PmlI-digested backbone (pPWQ74.3) using T4 DNA ligase. This vector is prepared from a bacterial expression vector, pTrcHis2-TOPO (Invitrogen, Carlsbad CA) by insertion of a fragment containing a ribosomal binding site followed by an initiation codon, a SpeI site, a PmlI site, and a

stop codon (Figure 8). Ligation products are transformed into the DH5α strain of *Escherichia coli* and selected on LB plates containing 100 mg L^{-1} ampicillin. The nucleotide sequences of the obtained constructs are confirmed by sequencing. Constructs with the correct sequence (pPWR59.2 for *Scopulariopsis* sp.; pPWR59.4 for *Penicillium canescens*) are then transformed into an *E. coli* strain from which the entire *gus* operon has been deleted (JEMA99.9). Transformants are selected on LB plates supplemented with 100 mg L^{-1} ampicillin, 40 mg L^{-1} isopropyl-β-D-thiogalactoside and 50 mg L^{-1} X-GlcA to induce expression of the cloned genes. A construct containing the gus gene of *E. coli* instead of a fungal gus gene (pPWR25.3) is used as positive control; the empty vector (pPWQ74.3) is used as negative control. As shown in Table 6, bacteria expressing either of the two fungal gus genes turn blue in the presence of the GUS substrate X-GlcA, while those containing the empty vector remain white.

Table 6. GUS activity of transgenic organisms without endogenous GUS activity that have been transformed with fungal gus genes

	Host	
Source of *gus* gene	***Escherichia coli***	**Leaves of rice plants**
Scopulariopsis sp.	+	+
Penicillium canescens	+	+
Escherichia coli	+	n.d.*
Staphylococcus sp.	n.d.*	+
Empty vector	0	n.d.*
Untransformed organism	n.d.*	0

GUS activity as visualized with X-GlcA added to the growth medium.

* Not determined.

[0149] For expression in plants, the two fungal gus genes are excised from the bacterial constructs by partial digestion with SpeI and PmlI restriction enzymes. Full-

length *gus* fragments are purified on a 1.2 % TAE agarose gel and extracted using the Qiagen Gel Extraction kit. Aliquots are then ligated to a plant expression vector from which the gus gene of a *Staphylococcus* species has been excised with SpeI and PmlI (pCAMBIA1305.2; Figure 9). This fuses the fungal gus genes to an upstream sequence comprising the GRP (glycine-rich protein) signal peptide and the catalase intron. The former mediates secretion in plant cells and the latter boosts expression levels in plants. A stop codon is located immediately downstream of the cloning site. Plasmid DNA of bacterial colonies obtained after transformation into DH5α and selection on LB plates containing 100 mg L^{-1} ampicillin is sequenced to confirm the cloning step.

[0150] Constructs with the correct sequence (pKKWA68.4 for *Scopulariopsis*; pPWT9.17 for *Penicillium canescens*), as well as the original pCAMBIA1305.2 construct (positive control), are transformed into a strain of *Agrobacterium tumefaciens* (EHA105) by electroporation (Hood et al., *Transgenic Res* **2**: 208-218, 1993; Sambrook *supra*). Transformants are selected on AB medium containing 50 μg mL^{-1} kanamycin (Chilton et al., *Proc. Natl. Acad. Sci. USA* 71: 3672-3676, 1974). Scutellum-derived callus of rice (*Oryza sativa* L. cv. Nipponbare or Millin) is then transformed with both constructs using the protocol of Hiei et al. (*Plant J* **6**: 271-282, 1994) and selecting for hygromycin-resistant plants.

[0151] Leaves of T0 and T1 plants are excised and incubated in a 50 mM sodium phosphate buffer (pH 7.0) containing 1mg mL^{-1} X-GlcA for up to 16 hours at 37°C (Jefferson, *Plant Mol Biol Reporter* **5**: 387-405, 1987). GUS activity is visible in plants transformed with both fungal gus gens or the *Staphylococcus* gus gene in as little as 1 hour as indicated by the presence of a blue precipitate in leaf tissue (Table 5; Figure 10A). No GUS activity (staining) is detected in leaves of untransformed plants. Leaf discs of plants transformed with pKKWA68.4 (*gus* from *Scopulariopsis*) and pPWT9.17 (*gus* from *Penicillium canescens*) stain the incubation medium significantly stronger than those of plants transformed with pCAMBIA1305.2 (gus from *Staphylococcus*) (Figure 10B). This suggests that secretion of fungal β-glucuronidases in plants is particularly efficient in plants.

Table 7: Identification of SEQ ID NOs.

SEQ ID NO: 1. DNA sequence of the *gus* gene of *Scopulariopsis* sp. isolate RP38.3
SEQ ID NO: 2. Amino acid sequence of GUS protein from *Scopulariopsis* sp. isolate RP38.3
SEQ ID NO: 3. DNA sequence of the *gus* gene of *Penicillium canescens* isolate RPK
SEQ ID NO: 4. Amino acid sequence of GUS protein from *Penicillium canescens* isolate RPK
SEQ ID NO: 5. DNA sequence of the *gus* gene of *Penicillium canescens* isolate DSM1215
SEQ ID NO: 6. Amino acid sequence of GUS protein from *Scopulariopsis* sp. isolate DSM1215
SEQ ID NO: 7. DNA sequence of the *gus* gene of *Gibberella zeae*
SEQ ID NO: 8. Amino acid sequence of GUS protein from *Gibberella zeae*
SEQ ID NO: 9. DNA sequence of the *gus* gene of *Aspergillus nidulans*
SEQ ID NO: 10. Amino acid sequence of GUS protein from *Aspergillus nidulans*
SEQ ID NO: 11. Amino acid sequence of GUS protein from *C. elegans*
SEQ ID NO: 12. Amino acid sequence of GUS protein from *D. melanogaster*
SEQ ID NO: 13. Amino acid sequence of GUS protein from *M. musculus*
SEQ ID NO: 14. Amino acid sequence of GUS protein from *R. norvegicus*
SEQ ID NO: 15. Amino acid sequence of GUS protein from *F. catus*
SEQ ID NO: 16. Amino acid sequence of GUS protein from *C. familiaris*
SEQ ID NO: 17. Amino acid sequence of GUS protein from *C. aethiops*
SEQ ID NO: 18. Amino acid sequence of GUS protein from *H. sapiens*
SEQ ID NO: 19. Amino acid sequence of GUS protein from *S. solfataricus*
SEQ ID NO: 20. Amino acid sequence of GUS protein from *T. maritima*
SEQ ID NO: 21. Amino acid sequence of GUS protein from *L. gasseri*
SEQ ID NO: 22. Amino acid sequence of GUS protein from *E. coli*
SEQ ID NO: 23. Amino acid sequence of GUS protein from *Staphylococcus* sp.
SEQ ID NO: 24. Primer ITS-fwd1
SEQ ID NO: 25. Primer ITS-rev4
SEQ ID NO: 26. Primer NS3
SEQ ID NO: 27. Primer NS6
SEQ ID NO: 28. ITS sequence from isolate RP38.3
SEQ ID NO: 29. ITS sequence from isolate RPK
SEQ ID NO: 30. 18S rRNA gene sequence from isolate RP38.3
SEQ ID NO: 31. 18S rRNA gene sequence from isolate RPK
SEQ ID NO: 32. Primer gus-fwd+T3
SEQ ID NO: 33. Primer gus-rev+T7
SEQ ID NO: 34. Primer gus(Scop)-fwd+SpeI
SEQ ID NO: 35. Primer gus(Scop)-rev+PmlI
SEQ ID NO: 36. Primer gus(Pcan)-fwd+SpeI
SEQ ID NO: 37. Primer gus(Pcan)-rev+PmlI

[0152] From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.